

A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: D81185
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1447 <TEF>
A:Cross-references: GB:AE002412; GB:AE002098; NID:97225783; PIDN:AAF40997.1; PID:9722579
A:Experimental source: serogroup B, strain MC58
A:Gene: NMB0569
C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
C:Keywords: oxidoreductase

Alignment scores:
Quality: 2289.00 Length: 447
Ratio: 5.121 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment block:
US-09-303-518D-125 x D81185 ..

Align seg 1/1 to: D81185 from: 1 to: 447

```
1 ATGATTAATCAATCAAAAGGCTTAACCTGCCCATCGCGGCAGACCGGA 50
|||||
1 Metilysiletylsylglyleuasnleuprollealaglyarprogl 17
51 GCAACCGCTTACGACGGCGCCGACATACGGAAGTCCGTTGGTGGCG 100
|||||
17 uGlnAlaValIyrAspGlyProAlaIleThrGluValAlaLeuLeuGlyG 34
101 AAGAATATGCGCGTATGCGCCCTGATGAAGATCAAGGAAGCGGATGCC 150
|||||
34 IuGlutYrAlaGlyMetArpProSerMetLysValLysGluGlyAspAla 50
151 GTCAAAAAGGCCAGTCTGTTGAACACAAAAGAAATCCGGCGTGT 200
|||||
51 ValLysLysGlyGlnValLeuPheGlnLysLysAsnProGlyValVa 67
201 GTTACTGCGCGGCTTCAGCAAAATCGCGCGATTCACCGTGGCGAA 250
|||||
67 IPhetrAlaProAlaSerGlyLysIleAlaIleIleHisArpGlyGlu 84
251 AGCGGCTACTGACGTGATGCTGATGCGCTGAAGGCAAGCAAGAAATC 300
|||||
84 YsArGValLeuGlnSerValValIleAlaValGluGlyAsnAspGluIle 100
301 GAGTTGAACGCTAGCAGCCTGGAAGCGTGGCAAACTTAAGCGCGAAGA 350
|||||
101 GlnPheGlnArGlyrAlaProGluAlaLeuAlaAsnLeuSerGlyGluG 117
351 AGTGGCGCGCAACCTGATCAATCGGCTTGGACGCGCTGGCGACCC 400
|||||
117 uValArGArGAsnLeuIleGlnSerGlyLeuTrpThrAlaLeuArGThrA 134
401 GTCCGTTGAGCAAAATTCCTCGCTGATGCCGAGCGCTTGGCATTTTC 450
|||||
134 rGProPheSerLysIleProAlaValAlaAspAlaGluProPheAlaIlePhe 150
451 GTCATATGATGAGACCAATCCGCTGGCTGGCGACCCCTACGGTCATTAT 500
|||||
151 ValAsnAlaMetAspThrAsnProLeuAlaAlaAspProThrValIleI 167
501 CAAGAAGACCGCGGAGATTCAACAGCGCGCTGTGTTATGGCGGTT 550
|||||
167 elysGluAlaIleGluAspPheLysArGglyLeuLeuValLeuSerArGL 184
551 TGACCGAAGCAAAATCCATGTTTGAAGGACGCTGGCGAGAGAGTCCG 600
|||||
184 eutThrGluArgLysIleHisValLysLysAlaIleGlyAlaAspValPro 200
601 TCTGAATAATGCTGCCAAACATCGAAACACATGAAATTGCGCGCGCATCC 650
```

```
|||||
201 SerGluAsnAlaIleAsnIleGluThrHisGluPheGlyLysProHispr 217
651 TGGCGGTTTGAAGGACGACACATTCATTCATCGACCGCGTGGCGCA 700
|||||
217 oAlaGlyLeuSerGlyThrHisIleHisPheIleGluProValGlyAla 234
701 ATAAACCGGTGACCATCAATTCACAGATGATTAATTCATGGCCGT 750
|||||
234 snLysThrValTrpThrIleAsnTrpGlnAspValIleThrIleGlyArG 250
751 TTCTTTCACACAGCGCGTTCGAACACCGACCGCGTATGCCCTAGTGG 800
|||||
251 LeuPheAlaThrGlyArgLeuAsnThrGluArgValIleAlaLeuGlyG 267
801 TTCTCAAGTCAACAACCGCGCTCTTGGTACCGCTTTGGGGCGCAAG 850
|||||
267 YSerGlnValAsnLysProArGLeuLeuArGThrValLeuGlyAlaLysV 284
851 TATCGCAATTAATCTGCGCGGCAATTCGATGACACAGACCGCGATTT 900
|||||
284 aISerGlnIleThrAlaIleGlyGluLeuValAspThrAspAsnArGValIle 300
901 TCCGCTTCGCTATTCAGCGCGGATTCACACAGCGCGCATTAATT 950
|||||
301 SerGlySerValLeuAsnGlyAlaIleThrGlnGlyAlaHisAspTyrLe 317
951 GGGAGGTACACCAATCAGATTCGCTATTCGAAGAAGCGCCGACGAAAG 1000
|||||
317 uGlyArGlyrYrHisAsnGlnIleSerValIleGluGluGlyArGSerLysG 334
1001 AGCTGTTCGGCTGGGTTGGCGCGCACCGACAAATATCTCATACGCGT 1050
|||||
334 IuLeuPheGlyrTrpValAlaProGlnProAspLysTrpSerIleThrArG 350
1051 ACAACCTCGGCGCATTTCTGAAAACAAACTCTTAAGTTCAACACAGC 1100
|||||
351 ThrThrLeuGlyHisPheLeuLysAsnLysLeuPheLysPheAsnThrAl 367
1101 CGTCAACGGGCGGCGGCGCGCATGTCGCGATGGTACTTCAGAGCGCG 1150
|||||
367 aValAsnGlyGlyAspArGAlaMetValProIleGlyrYrGluArGV 384
1151 TGATGCCCTTGGATATTCCTGCCACCCCTCTTGGCGGATTAATCGTC 1200
|||||
384 alMetProLeuAspIleLeuProThrLeuLeuLeuArGAspLeuIleVal 400
1201 GCGGATACGACAGCGCGGCGGATTCGCTGCTTGAATTCGACGAAAGA 1250
|||||
401 GlyAspThrAspSerAlaGlnAlaLeuGlyCysLeuGluLeuAspGlu 417
1251 AGACCTCGCTTGTGACGCTGTCGCGCGGCAATTCGAATACGGCGC 1300
|||||
417 uAspLeuAlaLeuCysSerPheValCysProGlyLysTrpGlyrYrGlyP 434
1301 CGCTGTTCGCAAAAGTGTGAACCATTCAGAAAGGAAGGC 1341
|||||
434 roLeuLeuArGlyValLeuGluThrIleGluLysGluGly 447
```

seq_name: p1r2:A81919

seq_documentation_block:

probable sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) chain A NM

C:Species: *Neisseria meningitidis*

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 16-Feb-2001

C:Accession: A81919

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491

A:Reference number: A81775; MUID:20222556

A:Accession: A81919

A:Status: Preliminary

A:Molecule type: DNA
A:Residues: 1-447 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CA84035.1; PID:g7379474
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: ngrA; NMA0752
C:Superfamily: Vibrio agnitoliticus sodium-translocating NADH dehydrogenase (ubiquinone)
C:Keywords: NAD; oxidoreductase

```
alignment_scores:
  Quality: 2251.00      Length: 447
  Ratio: 5.036          Gaps: 0
Percent Similarity: 100.000  Percent Identity: 97.763
```

```
alignment_block:
```

US-09-303-518D-125 x A81919

Align seg 1/1 to: A81919 from: 1 to: 447

1 ATGATTTAAATCAAAAGAGCTTAACCTGGCCATCGCGGGACAGCCGA 50
1 MettIleuylIleuysylsyeuInsneuProIlelaIagIaYpProgl 17
51 GCAGAGCGGTTTACAGACGGCGCGCCATTACCGAAGTCGCGTTCGTGGCG 100
17 uGInValIleIetyrAspGlyProValIleIethGluValAlaIeuleuSlyG 34
101 AAGATATATCCCGGATATGCGCCCTCGATGAAGAATCAAGAGGCGATGCC 150
34 IuGIuIyryIaIagIeIaYpProSerMetIySAllySgluGIyAspAla 50
151 GTCAAAAAGGCGCAGTGGTGTTTAAACAAAAAGAAATCCGGCGCGGT 200
51 ValIlySylSglIyAlaIleuIheGluAspIySylAsnProGIyAlaIva 67
201 GTTTACTGGCGCGGCTTCAGGCAAAATCGCGGCATTCACCGTGGCGAA 250
67 IPhethrIaIaProValSerGlyIySleIaIaIaIleIhIaSrGlyIyIuL 84
251 AGCGCGTACTTACAGTCAGTCGTGATGGTCCGTTGAAGCAGCAGCGAAATC 300
84 ySaIryAlaIeuInseryAlaIleAlaIaIaIuGIyAsnAspGluIle 100
301 GAGTTTGAACGGCTACACCTCAAGCGGTGGCAACTTAAACGGCGAGA 350
101 GluPhelGluIaIryrIaIaProGluAlaIeulaInsneuSerGIyGluGI 117
351 AGTGGCGCGCAACTGATCCATCCGGTTGTGACTGCGCTGGCGCACCC 400
117 uValaIaIrgAsnIleuIleGIInseryIyLeuIryPhAlaIeuaIryhIra 134
401 GTCCGTTACAGCAAAATCTCTGCGGTCGATGCCGAGCCGTTGGCATCTTC 450
134 rGrProPheSerIySleIProAlaIaIaAspAlaIaGluProPhAlaIlePhe 150
451 GTCAATGCGATGAGACACCAATCGGTGGCTGGCGACCTTAGCGATCAT 500
151 ValaInsIaIeMetAspThrAsnProIeulaIaIaAspProValIaIleI 167
501 CAAGAAGCGCGCGAGGATTTCAACGCGCGCGCTGTGGTATTAGCCGTT 550
167 eIySgluAlaIaIaGluAspPheAIdaIrgGlyIyLeuIeuaIleuSerAId 184
551 TGACCGAAGCGCAAAATCCATGTTTGTAAAGGCAAGCTGGCGGACAGCTCCG 600
184 euthrGluAIdaIySleIeIhIaIySylSylAlaIaIaGlyAlaAspValPro 200
601 TCTGAAAATCTGTCCAAATCAACAAACATGAATTCGGCGCGCGCATTC 650
201 SerGIuAsnAlaIaIaInsIleIeIuIrhIaIeIuIyGlyIyProhIsIpr 217
551 TGCGCGGTTGAAGTGGCGCGCACCATTCATTTCATCGACGCGGTCGGCGCA 700

217	calacIyIeusergelYthrhIsIlehsPhelleIuProValClYlhaI	234
701	ATAAACCGGTGGACATCAATTATCAAGATGATTAATTCACATTGGCCGT	750
234	snlysthrValtrPhrIrhIsantYrgInasrValIlealIleIglYarg	250
751	TTGTGTCACACAGCGCGTCTACACACGAGCGCGTGATGGCCCTAGGTGG	800
251	leuphealatrIglYargIeasnthrIglYarValIlealIealIeugIglY	267
801	TTTCGAATCAACAACCGCGCGCTGTGGCGATACCGGTTTGGGTGCGAAG	850
267	yserIlnalIasnlYsProalrIglelealrIghValIeulYalalYsv	284
851	TATCGCAATTAATCTGCGGGCGAATTTGGTTGTGACACAGACACCGCGTATT	900
284	alserIlnIlethrIaelYglueIeulalAsrIalalAsrIalrYalIle	300
901	TCGGTGGGTATGTGAACGCGCGCATTAACACAGCGCGCACGATTAATT	950
301	SerIglSerValIeasngIYalalIethrIngIYalIahIsarYtlle	317
951	GGGACGCTACACATCAATCGATTCCGTTATGCAAGAAAGCGCGACAAAG	1000
317	uGlYargIYrhIsasngIleIserValIleIgluIuGlYargSerIlysg	334
1001	AGCTTTGCGCGGGGTTGCGCGGACCGCGCAATATCCATACACGGT	1050
334	IuleIrhelYlTtrIalalalProIlnIProasrIlyIYserIlethrIarg	350
1051	ACAAACCTCGGCGCATTTCTCTATAAAACAAACTCTTCAAGTTCAACACAGC	1100
351	ThrIrhIeulIghIsrIheleulYasnlYserIeulrIheulrIheulrIhal	367
1101	CGTCAACGACGCGACGCGCGCCATGCTGCGCATTTGTAATTCACAGCGCG	1150
367	avalasngIglYasrIalrIameIvalrIeolYthrIYrgIuIargY	384
1151	TGATGCGCGTGGATTTCTGTGGCGACCGCGTGGTTTGGCGGATTTAAATGCTC	1200
384	almerIProIeulaspIleleuProIrhIleleulIeulalalAsrIeulIeal	400
1201	GGCGATACCGACAGCGCGCGACGCGCATTTGGGTGCTTGGAAATTCAGAGAGA	1250
401	GIYasrIrhIasrIasrIalIeIalIeulIYgIysIeulIeulalAsrIgluI	417
1251	AGACCTCGCTTTGTGACAGCTTCGTGCTGCCGCGGCAATATGCAATACGGCC	1300
417	uAsrIeulalIeulYsSerIheIalYsProIglYulYstIglYulYrgIY	434
1301	CGCTGTGCGGCAAACTGCTGGAAACCATTTAGAAAGGAGAGCG	1341
434	IoleIeulalrIglYsValIeulIupIrhIleIglIYlYsgIuIglY	447

seq_name: p1r2:I64002

seq_documentation_block:

sodium-translating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain H10164
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 02-Sep-2000 #text_change 02-Mar-2001
C:Accession: 164002; A64003
R: Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kierlavage
: Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodet, A.; Kelley, J.M.; Welsh
: D.M.; Brandon, R.C.; FINE, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Ventec
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: 164002
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-447 <TIGR>

A:Cross-references: GB:LA2023; TIGR:H10164; GB:U32702; NID:g1573118; PIDN:AAC21836.1; P1
 A:Note: the sequence is revised in GenBank entry U32702, PID:g1573122
 C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
 C:Keywords: oxidoreductase

alignment_scores:
 Quality: 1629.00 Length: 448
 Ratio: 4.093 Gaps: 2
 Percent Similarity: 88.839 Percent Identity: 68.973

alignment_block:
 US-09-303-518d-125 x 164002 ..

Align seg 1/1 to: 164002 from: 1 to: 447

```

1 ATGATTAATCAAAAAGGTCTAATCGCCATCGCGGCGAGACCGGA 50
1 MetIleThrIleLysGlyLeuAspLeuProIleAlaGlyLysProAl 17
51 GCAGCCGTTTACGAGCGCGCCCATTCAGCAAGTCGCGTGTGCGG 100
17 aGlnValIleHisSerGlyAsnAlaValAsnGlnValAlaIleLeuGly 34
101 AAGATATGCGCGGTATGCGCCCTCGATGAAAGTCAGGAAGCGGATCC 150
34 IuGluIleValGlyMetArgProSerMetLysValArgGlyLysPylVal 50
151 GTCAAAAAGGCGCAAGTGTCTTGAAGCAAAAAGATCGGGCGTGT 200
51 ValLysLysGlyGlnValLeuPheGlnLysPylAsnProGlyValAl 67
201 GTTACTGCGCGCGCTTCAGGCAAAATCGCGCGATTCACCGTGGCAAA 250
67 ePheThrAlaProAlaSerGlyThrIleThrAlaIleAsnArgGlyL 84
251 AGCGGCTACTTCAGTCACTGATGATGCGCGTGAAGGCAAGCAAGAAATC 300
84 YsArgValLeuGlnSerValValIleAsnValGlyLysPylLysIle 100
301 GAGTTGAAGCGTACGACACGCGTGAAGCGCGTGAAGCAAACTTAACGGCGAAGA 350
101 ThrPheAlaLysTyrSerThrGlnGlnLeuAsnThrLeuSerSerGly 117
351 AGTGGCGCGCAACCTGATCAATCGCGTGTGAGTCCGCTGGCAACC 400
117 nValLysGlnAsnLeuIleGlnSerGlyLeuThrPheAlaLeuArgH 134
401 GTCCGTTAGCAAAATTCCTGCGCTGCATGCGCGAGCGGTCGCATCTTC 450
134 rGpProPheSerLysValProSerIleGlnSerGlnAlaSerSerIlePhe 150
451 GTCAATGCGATGACACCAATCGCGTGGCGCGACCTCGGTCATAT 500
151 ValAsnAlaMetAspThrAsnProLeuAlaAspProSerValValle 167
501 CAAGAAGCCGCGAGATTTCAAGCGCGCGCTGTGATTTGAGCCGTT 550
167 uLysGlnLysSerGlnAspPheThrAsnGlyLeuThrValLeuSerArg 184
551 TGACCGAAGCAAA..ATCCATGTTTGAAGGAGAGTGGCGCGAGCGTG 597
184 eUpheProSerLysPylProLeuHisIleLeuLysValGlyAspSerAsnIle 200
598 CCGTCTGTAATGCTGCCAATCATGCAAAACATGATTCGGCGCGCGCA 647
201 ProThrAlaAspLeuGlnAsnLeuGlnIleHisAspPheThrGlyVal 217
648 TCGTCCGCGTTGAGTGCAGCAACATTCATTCATGAGCGCGTGGCGG 697
217 sProAlaGlyLeuValGlyThrHisIleHisPheIleAspProValGly 234
698 CGAATAAACCCTGTGAGCAATCATTAATGATGTAATTCACATTCG 747

```

```

234 legInLysThrValThrPheHisIleAsnTyrGlnAspValIleAlaValGly 250
748 CGTTGTTTGCACACAGCGCGCTGCAACACCGACCGCGTATTCCTAGG 797
251 LysLeuPheThrThrGlyGlnLeuLysTyrSerGlnArgValIleSerLeuAl 267
798 TGGTTCTCAAGTCAACAAACCGCGCTGTGCGTACCGTTTGGTGGCA 847
267 aGlyProGlnValLysGlnProArgLeuValAlaGlyThrIleGlyAla 284
848 AAGATTCGCAAAATTTACTGCGCGCAATGGTTGACACAGCAACCGCGT 897
284 snLeuSerGlnLeuThrGlnAsnGlnLeuSerAlaGlyLysAsnArgVal 300
898 ATTTCCGCTTCGATTTGAACGCGCGATTCACCAAGCGCGGACATTA 947
301 IleSerGlySerValLeuLysGlyGlnIleAlaLysAspSerHisAsp 317
948 TTTGGACCGCTACCAATCAGATTCGTTATCGAAGAAGCGCGACGA 997
317 rLeuGlyArgTyrAlaLeuGlnValSerValIleAlaGlyLysGln 334
998 AAGAGCTGTTCGCGTGGTTCGCGCGACCGCAACAAATTCATCATCAG 1047
334 YsGlnPhePheGlyTyrIleMetProGlnAlaAsnLysTyrSerValThr 350
1048 CGTCAAAACCTCGCGCATTCCTGAAAAAACCTTCAGTTGACAGC 1097
351 ArgThrValLeuGlnIleHisPheSerLys...LysLeuPheAsnIleThr 366
1098 AGCGGTCAACGCGCGCGACCGCGCATGCGCGATTCGATTCAGAC 1147
366 rSerGlnAsnGlyGlyGlnArgAlaMetValProIleGlySerTyrGln 383
1148 GCGTGAATGCCCTTGATATCTGCGCCACCGCTGTTGGCGCATTAATC 1197
383 rValMetProLeuAspIleLeuProThrLeuLeuArgPylLeu 399
1198 GTGGCGGATACGACAGCGCGGATTCGTTGATTTGGAATGGACGA 1247
400 ValGlyAspThrAspGlyAlaGlnLeuGlnLysGlyLeuGlnLeuAsp 416
1248 AGAAGACCTCGCTTGTGACGCTGCTGCGCGGCAAAATGCAATGCG 1297
416 uGlnAspLeuAlaLeuLysSerPheValCysProGlyLysTyrGlyTyrG 433
1298 GCCCGCTGTTCGCAAAAGTGTGAAACCATTCAGGAAGGAGGC 1341
433 YsSerIleLeuArgGlnValLeuAspLysIleGlyLysGlnGly 447

```

seq_name: p1r2:AG0393

seq_documentation_block:

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain A [imported] - Yersinia pestis (st
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AG0393
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586560
 A:Accession: AG0393
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-447 <KDR>
 A:Cross-references: GB:AL590842; PIDN:CA092475.1; PID:g15981176; GSPDB:GN00175
 C:Genetics:
 A:Gene: ngrA
 C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino
 C:Keywords: oxidoreductase

alignment_scores:
 Quality: 1596.00 Length: 447
 Ratio: 4.082 Gaps: 0
 Percent Similarity: 87.472 Percent Identity: 68.680

alignment_block:
 US-09-303-518D-125 x AG0393 ..

Align seg 1/1 to: AG0393 from: 1 to: 447

```

1 ATGATTAATCAAAAAAGCTTAACCTGCGCGGCGGACCGGCA 50
|||||
1 Metilelystlelysglyleuasproulealaglyalaprova 17
17 IGLnThrllleGlnasrglyproalalehshisvalalaleuengl 34
51 GCAGGCCCTTACGACGGCGCGCCATTACCGAAGTGGCGTGGCGG 100
|||||
101 AAGAAATAGCGGATAGCGCGCCCTCGATGAAGACGAGGAGGATGCC 150
|||||
34 IuglutyValglYmetarprosermetleuValGlnGluGlyaspGln 50
151 GTCAAAAAGGCGAAGTGTGTTGAAGACAAAAGAAATCGCGCGGTGT 200
|||||
51 VallyslsglyGlnalaleuPheGlnasplyslsasnproGlyValle 67
201 GTTTACTCGCGCGCTTCAGGCAAAATGCGCGGATTCACCGTGGCGAA 250
|||||
67 upherthralproalaserGlylsIleseralalleanaraglyGlna 84
251 AGCGGTACTCTACGTCAGTGGTGGATGGCGGTAAGCGAAGCAAGAA 300
|||||
84 rgaValleuGlnserValIleGluValIleGluGlyaspGlnIle 100
301 GAGTTTGAACGCTACGACCTGAAGCGCGTGGCAAACTTAACGCGGAGA 350
|||||
101 ProPheGlnhshstYrAlalaglnGluLeuasnGlnleuSerAaspGln 117
351 AGTGGCGCGCAACCTGATCCAACTCGGTTGTGACTGGCGTGGCGACC 400
|||||
117 nValGlnhshshleuLeuThrSerglyLeuThrPhralaleuarglyTra 134
401 GTTCGCTGACAAATTCCTGCGCGTGCATGGCGAGCGTGGCGCATCTTC 450
|||||
134 rgrProPheSerlyThrProValIProGlySerArgProArgAlalPhe 150
451 GTCAATGGGATGACCAATCCGCTGGCTGGCGACCCCTAGCGTCATTAT 500
|||||
151 IleseralIleMetAspThrGlnProleuAlalAspProGlnValIle 167
501 CAAAGAGCGCGCGAGATTTCAACGCGCGCTGTGATTATGACCGCTT 550
|||||
167 ealathrgIuSerGlnalalPheasnhsGlyleuThrValleuThrArg 184
551 TGAACGAGCGCAAAATTCATGTTTGAAGCGCGTGGCGAGCGAGCGCG 600
|||||
184 eulthAspIlyslValhshshValCysHshshAlalalGlyAlalValhsh 200
601 TGTGAATGTCTCCACATCGCAACACATGAATTCGGCGCGCGCATCC 650
|||||
201 ArgHshstHAspThrGlnValhshhshshshshshshshshshshsh 217
651 TGGCGGTTTGAAGTGGCGACGACATTCATTCATCGACGCGTGGCGGGA 700
|||||
217 oalaglyleuValIglYthrHshshshshshshshshshshshshsh 234
701 ATGAACGCTGTGAGCAATCAATTAATGAATGATTAATCAATTCGCGGT 750
|||||
234 hrlYsMetValItrPhshshshshshshshshshshshshshshshsh 250
751 TTGTTTGAACAGCGCGTCTGAACACGAGCGCGTGAATTCGCGCTAGTGG 800

```

```

|||||
251 LeuPheThrArgGlyGluLeuCysThrAspArgIleValAlalaleuAlaG 267
801 TTCTCAAGTCACAAACCGCGCTCTTGGTACCGCTTTGGGTGGCGAAG 850
|||||
267 rProGlnValasnGlnProIleleuLeuArgThrArgleuGlnIleAsp 284
851 TATCGCAATTTACTGCGCGGATGGTGTGACACAGCAACACCGCTGAT 900
|||||
284 euserGluLeuThrAlaGlylsleuGlysglnGlyAspAsnArgIlele 300
901 TCCGTTTCGGATTGACAGCGCGCATTTACACAAGCGCGCGCATTTAT 950
|||||
301 SerGlySerValleuSerGlyThrAlaPheSerAlaThrshshshshsh 317
951 GGGACGCTACACATCAATTCGCTTTCGACAGAGCGCGCGCGCAAG 1000
|||||
317 uGlyArgPhehshshshshshshshshshshshshshshshshshsh 334
1001 AGCTGTTCGGGCTGTGGCGCGACCGGACCAATATCTCATCGCGGT 1050
|||||
334 lueuPheGlyTrpValMetProGlyArgAspIlyslYserIleThrArg 350
1051 ACAACGCTCGCGCATTCCTGAAAACAAACCTTCATCACTCAACACAG 1100
|||||
351 ThrThrleuGlnhshshshshshshshshshshshshshshshshsh 367
1101 CGTCACAGCGCGCGCGCGCGCGCATTCGCGGATGGTACTTACGACGG 1150
|||||
367 pMetHshshshshshshshshshshshshshshshshshshshshsh 384
1151 TGATGCCCTTGATATCTCCACCGCTGTTGGCGGATTTAATCTGTC 1200
|||||
384 alMetProleuAspIleleuAlaThrshshshshshshshshshshsh 400
1201 GCGGATACGACAGCGCGCGCGCATTCGCTTGGATTTGACAGAA 1250
|||||
401 GlyAspThrAspSerAlaGlnAlalaleuGlyCysleuGlnleuAsp 417
1251 AGAGCTTCGCTTGGCGCGCTGCTGGCGCGCGCAATGCAATGACGCG 1300
|||||
417 uAspLeuAlaleuCysThrPheValCysProGlyslYtrGluTyrGly 434
1301 CGCTGTTCGCGCAAGTCTGGAACCAATTCGAGAGGAGGAGC 1341
|||||
434 rovalleuArgAspIleleuThrGlnIleGlnGlnGlnGly 447

```

seq_name: p1r2:S51015

seq_documentation_block:
 sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain ngra [V
 N:Alternate names: ngra protein
 C:Species: Vibrio alginolyticus
 C:date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 16-Feb-2001
 C:accession: S51015; S66365
 R:Beattie, P.; Tan, K.; Bourne, R.M.; Leach, D.; Rich, P.R.; Ward, F.B.
 FEBS Lett. 356, 333-338, 1994
 A:title: Cloning and sequencing of four structural genes for the Na(+)-translocating
 A:reference number: S51013; MUID:9510445
 A:accession: S51015
 A:molecule type: DNA
 A:residues: 1-446 <BEA>
 A:cross-references: EMBL:Z37111; NID:g663268; PIDN:CAA85476.1; PID:g663270
 A:accession: S66365
 A:molecule type: protein
 A:residues: 1-10 <BBM>
 C:genetics:
 A:gene: ngra
 C:superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino
 C:keywords: FMN; oxidoreductase

alignment_scores:

Quality: 1468.50 Length: 447
Ratio: 3.814 Gaps: 1
Percent Similarity: 86.130 Percent Identity: 62.416

Alignment block:

US-09-303-518D-125 x SS1015 ..

Align seg 1/1 to: SS1015 from: 1 to: 446

```

1 ATGATTAATAAATAAAGCTGTAACCTGATCCGATCGGCGGACAGACCGGA 50
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 MetIethrIleIysIleGlyLeuAspLeuProIleIleIleGlyThrProse 17
51 GCACCCGTTTACGACGGCCGCGCATTTACCAAGTCGCGTGGTGGCG 100
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 IclnValIleAsnAspGlyLysThrIleIleIleValIleLeuLeuGlyG 34
101 AAGATATGCGGATGTCGCGCCCTCGATGAAGAATCAAGAGGAGGATGCC 150
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34 IuGIuTyValGIuMetIarProThrMetHisValIarGValGIuAspIu 50
151 GTCAAAAAGGCCAAGTGTCTTGAAGACAAAAGAAATCCGGCGGTGCT 200
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 ValIysIysAlaIleValIleuPheGluAspIysIysAsnProGIuVally 67
201 GTTACTCGCGCGCTTCAGCAAAATCGCGCATTCACCGTGGCGGAAA 250
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 sPheThrIarProIleIleIleGlyLysValIleGIuValIleAsnArgIyAla 84
251 AGCGGCTCTCAGTCAGTGTGATGCTGTCGCGTGAAGCAACGACGAAATC 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84 ysaIyValIleuGIuSerValIleGIuValIleGIuValIleGIuGIuVal 100
301 GAGTTTGAACGCTAGCGACCTGAAGCGCTGCAACTTAAGCGCGCAAGA 350
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 ThrPheAspIysPheGIuAlaIleIleuSerGIuLeuAsnArgIuVal 117
351 AGTGGCGCGCAACTGATCCAAATCGGTTTGGACTGCGCGTGGCGACCC 400
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117 ILeuIysThrGIuLeuValIleAspSerGIuLeuThrIleAlaLeuArgIh 134
401 GTCCGTTACAGAAATTCCTGCGCTGATGCTGCGCGCGTTCGACATTC 450
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 rgrProPheSerIysValIleProIleIleGIuSerThrIlySAlIlePhe 150
451 GTCAATCGATGACACCAATCCGCTGCTGCGACCTTACGGTCAATTAT 500
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 ValThrIleMetAspThrIarProIleuAlaIleIysProGIuLeuIleI 167
501 CAAGAAGCGCGCGGATTTCAACCGCGCTGCTGCTTGAACCGCTT 550
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 eAsnGIuGIuGIuGIuAlaPheIleIleGIuAspIleIleuSerAlaI 184
551 TGACGGAACGCAAAATCATGTTTGAAGCAGCTGCGCGACGACGCGCG 600
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 eutHrGIuGIuLysValIyValCysIys...SerGIuThrIerLeuPro 199
601 TCTGAAAATCTGCGCAACATGGAACACATGAATTCGGCGCGCGCATCC 650
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
200 ArgSerSerGIuSerAsnValGIuGIuHisValIleAspIyProHisPr 216
651 TGCCGCTTGAAGCGACGACATTCATTTTCATCGACGCGTGGCGGCA 700
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 oAlaGIuLeuAlaGIuThrHisMetHisPheLeuTyProValAsnAlaG 233
701 ATAAAACCGTGTGACCATCAATTAATCAAGATTAATTAATTCAGCGCT 750
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
233 IuAsnValAlaIarPserIleAsnTyGIuAspValIleAlaPheGIuLys 249
751 TTGTTTGAACAGAGCGCTGGAACACCGAGCGGTGATTCCTTAGGTGG 800
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 LeuPheLeuThrGIuLeuLeuTyThrAspArgValIleSerIleuAlaGI 266

```

```

801 TTCTCAAGTCACAAACCGCGCTCTGCTACGCTTTGGTGGCGGAAG 850
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
266 yProValValAsnAsnProArgLeuValArgThrValIleGIuAlaSerL 283
851 TATCGCAAAATTTACTGCGGGAATGTTGATACACAGCAACCGCTGAT 900
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 euAspAspLeuThrAspAsnGIuLeuMetProGIuValAlaArgValIle 299
901 TCCGGTTCGCTATTGAACGCGCGGATTTACACAAGCGCGCGCATTTT 950
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 SerGIuSerValIleuThrGIuThrHisAlaIleGIuProHisAlaIle 316
951 GCGACCTTACACATCAATTCCTTATCGAAGAAGCGCGGACGGAAG 1000
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
316 uGIuArgTyHisGIuGIuValSerValIleuArgGIuGIuGIuGIuG 333
1001 AGCTGTTCGCTGGCTGTTGCGCGACCGGACAAATATCTCATCAGCGT 1050
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
333 IuLeuPheGIuTrpAlaMetProGIuLysAsnIysPheSerValIleArg 349
1051 ACACCCCTGCGCATTCCTCGAAGAAACAATCTTCAAGTTCAACAGAC 1100
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
350 SerPheLeuGIuHisValIlePheIysGIuIleuPheAsnMetThrTh 366
1101 CGTCAACGCGCGGACCGCGCATGTCGATGTCGATTCGACGCGG 1150
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
366 rThrAsnGIySerAspArgSerMetValProIleGIuAsnTyGIuArg 383
1151 TGATGCGCTTGATATCTGCGCCACCTGCTTTGCGCGATTTAATCGCT 1200
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
383 alMetProLeuAspMetGIuProThrIleuLeuArgAspLeuCysAla 399
1201 GCGATACGACAGACGCGGACGATTTGGTTCCTTGAATTCGACAGAGA 1250
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
400 GIyAspThrAspSerAlaIleIleuGIuAlaIleuGIuLeuAspGIuGI 416
1251 AGACCTCGCTTGTGCAGCTTCCTGCGCGGCAAAATGCAATTCGCGCC 1300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
416 uAspLeuAlaIleuCysThrPheValCysProGIuLysTyGIuTyGIu 433
1301 CGCTGTGCGCAAACTGCTGCAAAACCATTTGACAGGAAGCC 1341
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
433 hrLeuLeuArgGIuLysLeuAspThrIleGIuLysGIuGIu 446

```

seq_name: p1r2:G82094

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) nqrA VC2295 [import

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 16-Feb-2001

C:Accession: G82094

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

chardson, D.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: G82094

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1462 <HEI>

A:Cross-references: GB:AE004300; GB:AE003852; NID:99656850; PIDN:AAF95439.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Map position: 1

C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino

C:Keywords: oxidoreductase

alignment_scores:
Quality: 1455.50 Length: 447
Ratio: 3.810 Gaps: 1

Percent Similarity: 85.459 Percent Identity: 61.074

Alignment_block:

US-09-303-518D-125 x G82094

Align seg 1/1 to: G82094 from: 1 to: 462

```

1 ATGATTAATAAAGGCTTAACCTGCCATGCGGCGACGACCGA 50
   |||||
17 MetIleThrIleLysGlyLeuAspLeuProIleAlaGlyThrPro 33
51 GCAAGCCGTTACGACGCGCGCATACGACGCGCTGCTGGCG 100
   |||||
33 rGlnValIleSerAspGlyValIleLysValAlaLeuLeuGly 50
101 AAGATATGCGGTATGCGCCCTCGATGAAGTCAAGGAGCGATGCC 150
   |||||
50 IuGIuTyValGlyMetArgProThrMetHisValArgValGlyAspGlu 66
151 GTCAAAAAAGCCAAAGTCTGTTGAAGACAAAAAGAAATCCGGCGCT 200
   |||||
67 ValIysLysAlaGlnIleLeuPheGluAspLysAsnProGlyValIy 83
201 GTTTACTGCGCGCGCTTACGCAAAATGCGCGCGATTCACCGTGGCGAAA 250
   |||||
83 sPheThrSerProValSerGlyLysValValGluIleAsnArgGlyAla 100
251 AGCGGCTACTTCAGTCACTGCTGATGCGGTGAAGCAACGACGAAATC 300
   |||||
100 yArgValIleuGlnSerValIleGlyValAlaGlyAspAspGlyValI 116
301 GAGTTTGAACGCTACGCACTGAGCGCGTGAACCTTAACGCGCGAGA 350
   |||||
117 ThrPheAspLysPheGluAlaAsnGlnLeuAlaSerLeuAsnArgAspAl 133
351 AGTGGCGCCGACCTGATCCATCCGCTTGTGACTGCGCTGGCGACCC 400
   |||||
133 AlIeLysThrGlnLeuValGlnSerGlyLeuThrPheAlaPheArgThrA 150
401 GTCCGCTTGACAAATTCCTCGCGTGCATGCCAGCGCTTGCCATCTTC 450
   |||||
150 rGProPheSerLysValProAlaIleAspSerThrSerGlyAlaIlePhe 166
451 GTCAATGCCATGACACCAATCCGCTGGCTGGCGACCTAGGCATATAT 500
   |||||
167 ValThrAlaMetAspThrAsnProLeuAlaAlaGluProIleValIle 183
501 CAAAGAGCGCGCGAGATTCAAACGCGCGCTTGGTATTTGAAGCCGTT 550
   |||||
183 eaGlnGlnGlnSerGluAlaPheValAlaGlyLeuAspValLeuSerAla 200
551 TGACCGACGCAAAATTCATGTTTGTGAAGCAGCTGGCGCGACGCTGCC 600
   |||||
200 euhThrGlyLysValIyValIyValIy...LysGlyThrSerLeuPro 215
601 TCTGAATAAGCTCCCAACATCGAACAACATGATTCGGCGCGCGCATCC 650
   |||||
216 ArgSerGlnGlnProAsnValGlnGlnHisValPheAspGlyProHisP 232
651 TCGCGGTTTGAAGTGCACGACATTCATTCATTCAGCGCGCTGGCGCGA 700
   |||||
232 oAlaGlyLeuAlaGlyThrHisMetHisPheLeuTyProValSerAla 249
701 ATAAACCGTGTGACCATCATTAATTCAGATGATTAATTCATTCGCGCG 750
   |||||
249 sPheIleValIleThrSerIleAsnTyGlnAspValIleAlaValGly 265
751 TTGTTTCAACAGCGCGCTGTAACAGCGAGCGGATGATGCGCTAGGTG 800
   |||||
266 LeuPheLeuThrGlyIleuLeuTyThrGlnArgValValSerLeuAlaG 282
801 TTCTCAAGTCACAAACCGCGCTTTCGCTACCGTTTGGTGGCAAG 850
   |||||

```

```

282 yProValValAsnLysProArgLeuValArgThrValMetGlyAlaSerL 299
851 TATCGCAATTTACTGCGCGGATTTGGTTGACACAGACAAACCGCTGAT 900
   |||||
299 euGIuGlnLeuValAspSerGlnIleMetCProGlyValValArgIleI 315
901 TCCGTTTCGATTGAACGCGCGGATTTACACAAGCGCGCATTAATTT 950
   |||||
316 SerGlySerValIleuSerGlyThrIlySalAthrGlyProHisAlaTy 332
951 GGAACGCTTACCAACATGATATTCCTGTTATTCAGAAAGCGCGAGCAAG 1000
   |||||
332 uGlyArgTyHisLeuGlnIleValSerValIleuArgGlnGlyArgAsp 349
1001 AGCTGTTCGGCTGGTTCGCGCGCGACGCGCAAAATATCTCATACAGCG 1050
   |||||
349 IuLeuPheGlyTyPAlaIleMetProGlyLysAsnLysPheSerValThr 365
1051 ACAACCCCTCGGCCATTTCTGTAACAAACAACTTCAGATTCAACACAG 1100
   |||||
366 SerPheLeuGlnHisLeuPheLysGlyGlnValTyAsnMetThrThr 382
1101 CGTCAACGCGCGCGCGCGCATGCTGCGATTTGCTACTTACGACGCG 1150
   |||||
382 rThrAsnGlySerAspArgSerMetValProIleGlyAsnTyGlnLys 399
1151 TGATGCCCTTGATATCTCGCCACCTGCTTGTGCGGATTAATTCGTC 1200
   |||||
399 AlMetProLeuAspMetClnProThrIleuLeuArgAspLeuGlyAla 415
1201 GCGGATACCGACAGCGCGACGATTTGGGTTGCTTGAATTTGACGAA 1250
   |||||
416 GlyAspSerAspSerAlaValAlaArgLeuGlyAlaLeuGlnLeuAsp 432
1251 AAGACTCGCTTTGCTGACCTTCGCTGCGCGCGCAATTAACGATACG 1300
   |||||
432 uAspLeuAlaLeuGlyCysThrPheValCysProGlyLysTyGlnTy 449
1301 CGCTGTTCGCAAAAGTCTGGAACCAATTCGAAGGAGGAGG 1341
   |||||
449 IuLeuLeuArgGlnIuGlyLysLeuAspLysIleGlnLysGlnGly 462

```

seq_name: pir2:H83272

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) nqr1 chain PA2999 [C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 16-Feb-2001
 C:Accession: H83272
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lorry, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A:Reference number: A82950; M01D:2043737
 A:Accession: H83272
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-445 <ST0>
 A:Cross-references: GB:AE004724; GB:AE004091; NID:99949083; PIDN:AA06387.1; GSPDB:GN
 A:Experimental source: Strain PA01
 C:Genetics:
 A:Gene: nqrA; PA2999
 C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
 C:Keywords: oxidoreductase

alignment_scores:

Quality: 1353.00 Length: 447
 Ratio: 3.627 Gaps: 1
 Percent Similarity: 83.445 Percent Identity: 58.613

alignment_block:


```

3 11ethrValasnaIrgGlyLeuAspLeuSerLeuGInglySerProLySG1 19
51 GCAAGCCGTTTACGAC.....GGCCCGGCCATTACCGAAGTCGGTGGC 94
19 userglyPheThrValAsnLysIleAspProGluPheValSerIleAspLeu 35
95 TTGGCGAAGATATGCGCGTATGCGCCCGCCCGATGCAAGTCAAGAGAGC 144
36 ....ArgProPheGlnProLeuSerLeuLysLeuLysValGIngly 50
145 GATGCCGTCAAAAAAGGCCAAGTGGTGTGTAAGCAACAAAAGATCCGGG 194
51 AspaIValLysSerGlyAlaProIleAlaGluTyrLysHisPheProAs 67
195 CGTGGTGTACTGCGCGCGCTTACGCAAAATGCGCGCGATTCACCGTG 244
67 nThrTyrIleThrSerHisValSerGlyValAlaThrAlaIleArgArg 84
245 GCGAAAGCGCGCTTACTTACGTCAGTCAGTCGAT...GCCGTGAAGCAAC 291
84 LysnLysArgSerLeuLeuAspValIleIleLysLysThrProGlyPro 100
292 GACGAAATCGAGTTTGAACGCTACGACACCGTACGCGCGTGAACCTAAG 341
101 ThrSerThrGluTyr.....ThrTyrAspLeuGlnThrLeu 113
342 CGCGCAAGACAGTGGCGCGCAACATGATCCATTCGGTTTGGACTGCGC 391
113 rArgSerAspLeuSerGlnIlePheLysGlnAsnGlyLeuPheAlaLeu 130
392 TGCGACCGCGTGGTTCAGCAAAATTCCTGCGCGTGGTGGTGGCAG...CCG 438
130 leuLysGlnAlaProPheAsp...IleProAlaIleProIleGlnThrPro 145
439 TTGCGCATTCCTGCAATGCGATGACACCAATCCGCTGGTGGTGGCGCC 488
146 ArgAspValPheIleAsnLeuAlaAspAsnArgProPheThrProSerPr 162
489 TACGCTGATATC.....AAGAAGCGCGCGAGAT 520
162 ogLysHisIleAlaLeuPheSerSerArgGlnGluGlyValr 179
521 TCAACACCGCGCTGGTGTATGCGCGTGGTGAACCGCAACCAATCCAT 570
179 heValValGlyValArgAlaIleAlaLysLeuPheGluLeuArgProHis 195
571 GTTGTGAAGCGACGCGCGAGACGTGCGCTGAAATGCTGCAACAT 620
196 IleValPheArgAspArgLeuThrLeuProThrGlnGluLeuLysThrI 212
621 C...GAACACATGATGCGCGCGCGCGCGCGCTGCGGTGGAGTGGCA 667
212 eAlaHisIleLeuHisThrValSerGlyProPheProSerGlySerProSerI 229
668 CGCACATTCATTCATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
229 leHisIleHisSerValAlaProIleThrAsnGluLysGluValAlaPhe 245
715 ACATGATATGATGACATGATATTCATTCATTCATTCATTCATTCATTC 764
246 ThrLeuSerPheGlnAspValLeuThrIleGlnLysLeuPheLeuLysG 262
765 CGCTGCAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 814
262 yArgIleLeuHisGlnGlnValThrAlaLeuAlaGluTyrAlaLysLys 279
815 AACCG.....CGCGTGTGCGTACCGCGTGTGGTGGCGAAATGATCGGA 858
279 eSerLeuAlaArgTyrValIleThrThrLysGlyAlaSerPheSer 295
859 ...ATTACTGCGCGCGAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 905
296 leuIleAsnLeuAsnAspIleSerAspAsnAspThr...LeuIleSerG 311

```

```

906 TTGGTATTGCAACGCGCGATTACACAGGCGCGCACGAT...TATTGG 952
311 yAspProLeuThrGlyArgLeuLysLysGlnGluGluProPheLeuG 328
953 GACGCTACCAATTCAGATTTCGTTATGCAAGACGCGCGCAAGAG 1002
328 LysPheArgAspHisSerIleSerValLeuHisAsnProThrLysArgG 344
1003 CTGTGCGGTGGGTGGTGGCGCGCGCGCAAAATATCCATGACGCGTAC 1052
345 LeuPheSerPheLeuAlaGlyIleGlyPheAsnLysProThrPheThrLys 361
1053 AACCTTCGCGCATTCCTGTAACAAACAACTCTTCAAGTTCAAC.....A 1096
361 rTyrLeuSerGlyPhePheLysLysArgThrTyrThrAsnProAsp 378
1097 CAGCGCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1146
378 hAsnLeuHisGlyGluThrArgProIleIleAspThrAspIleTyrAsp 394
1147 CGCGTATCCCTTGGATATCCTGCGCACCGCTTGTGGCGGATTTAAT 1196
395 LysValMetProMetArgIleProValValProLeuIleLysAlaValI 411
1197 CGTGGCGCATTACCGACAGCGCGCGCGCGCATTTGGTTCGATTTGACG 1246
411 eThrLysAsnPheAspLeuAlaAsnGluLeuLysLeuGluValLysG 428
1247 AAGAAGACCTCGCTTGTGACGCTTCGTCGCGCGCGCGCAATACGATAC 1296
428 LysLysPheAlaLeuProThrIleLeuAspProSerLysThrGlnMet 444
1297 GCGCGCGCTTTCGCAAAAGTCTGGAACCATTGAAAGAA 1338
445 leuThrIleValLysGluSerLeuIleGluTyrAlaLysGln 458

```

seq_name: p1r2:E72040

seq_documentation_block:
 probable sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 16-Feb-2001
 C:Accession: E72040; G81623
 R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, Nature Genet. 21: 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; M01D:99206606
 A:Accession: E72040
 A:Molecule type: DNA
 A:Residues: 1-467 <RNA>
 A:Cross-references: GB:AE001656; GB:AE001363; NID:g4377047; PIDN:AD18882.1; PID:g437
 A:Experimental source: strain GW1029
 R:Read, T.D.; Brinham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis Morn and Chlamydia pneumoniae AK39
 A:Reference number: AB1500; M01D:20150255
 A:Accession: G81623
 A:Molecule type: DNA
 A:Residues: 1-467 <REA>
 A:Cross-references: GB:AE002164; GB:AE002161; NID:g7188939; PIDN:AAF37899.1; PID:g718
 A:Experimental source: strain AK39, HL cells
 C:Genetics:
 A:Gene: ngrA; CP00002
 C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
 C:Keywords: oxidoreductase

alignment_scores:

Quality:	464.50	Length:	464
Ratio:	1.585	Gaps:	15
Percent Similarity:	63.147	Percent Identity:	30.172

alignment_block:
US-09-303-518D-125 x E72040

Align seg 1/1 to: E72040 from: 1 to: 467

```

4 ATTAATCAAAAAAGGTAAACCTCCATCGCGGCGAGACCG...CA 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3 IleThrValAsnArgGlyLeuAspLeuSerLeuInglYserProLys 19
51 GCAACCCCTTACGAC.....GCCCGGCGCATTAACGAGTGGCGCTGC 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
19 uSerGlyPheThrAsnLysIleAspProGluPheValSerIleAspLeu 35
95 TTGGCGAAGAAATATGCGCGGTATGCGCCCTCGATGAAGATCAAGAGCG 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
36 ....ArgProPheGlnProLeuSerLeuLysValGluInglY 50
145 GATGCCGCTCAAAAAAGCCAAAGTGTGTTTAAACAAAAAAGATCCGGG 194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 AspAlaValLysSerGlyAlaProIleAlaGluThrLysHisPheProAs 67
195 GGTGGGTATTAAGTGGCGGCTTACGCAAAATGCGCGCATTCACCGTG 244
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 mThrTyrllePheSerHisValSerGlyValAlaThrAlaIleArgPrg 84
245 GCGAAAAAGCGCGTACTTCAAGTCAAGTGTAT...GCCGTGAAGGCAAC 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 LysHisLysArgSerLeuLeuAspValIleIleLysLysThrProGlyPro 100
292 GACGAATATGAGTTTGAAACGCTACCGACCTGAAGCGCTGGCAACTTAA 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 ThrSerThrGluTyr.....ThrTyrlAspLeuGlnThrLeu 113
342 CGCGAAGAGAGTGGCGGCAACCTGATCCAAATCCGTTTGATGATGCGC 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 ArgSerArgPheLeuSerGluIlePheLysGluAsnGlyLeuPheAlaLeu 130
392 TGGCAACCGCGCTGTTGACGAAATTCCTCGCGGTGATGCCGAG...CCG 438
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
130 LeuLysGlnArgProPheAsp...IleProAlaIleProThrGlnThrPro 145
439 TTGGCATATGCTGTCATGATGCGACACCAATCCGCTGGCGGCGACCC 488
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 ArgAspValPheIleAsnLeuAlaAspAsnArgProPheThrProSerPr 162
489 TACGGTCATTATC.....AAGAGCGCGCGAGGATT 520
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
162 ArgLysHisLeuAlaLeuPheSerSerArgGlnGluGlyPheTyrlP 179
521 TCAAAACCGCGCTGTTGATGACGCGCTTGAACGCAAAATCCAT 570
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
179 heValValGlyValArgAlaIleAlaLysLeuPheGlyLeuArgProHis 195
571 GTTGTGAAGCAGCTGGCGGAGAGTGGCTGAAATATGCTGCACAAAT 620
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196 IleValPheArgAspArgLeuThrLeuProThrGlnGluLeuLysThrI 212
621 C...GAACACATGAATTCGCGCGCGCGCATCTCGCGGTGATGGAGGA 667
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212 eAlaHisLeuHisThrValSerGlyProPheProGlySerProSerPr 229
668 CGGCAATTCATTCATGAGCGCGTGGCGGCGGCAATTA...ACCGTGG 714
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
229 LeuHisLeuHisSerValAlaProIleThrAsnGlnLysGluValAlaPhe 245
715 ACATCAATTAACAAGATGAATTAACATGGCGGCTGTTGCAACAGG 764
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
246 ThrLeuSerPheGlnAspValLeuThrIleGlyHisLeuPheLeuLysG 262
765 CCGTGTGAACACGAGCGCGTGAATGCGCGTGAAGTGTGTCGAAGTAA 814
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
262 yArgIleLeuHisGlnGlnValThrAlaLeuAlaGlyThrAlaLeuLys 279

```

```

815 AACCG.....CGCTCTGCTACCGCTTTGGGTGCGAAAGTATCGCA 858
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
279 erSerLeuArgArgTyrlValIleThrLysGlyAlaSerPheSer 295
859 ...ATTACTGGGCGGAGATGTTGATGACACAGACACCGCGTATTCGG 905
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
296 LeuIleAsnLeuAsnAspIleSerAspAsnAspThr...LeuIleSerG 311
906 TTGGTATTAAGCGCGCGATTCACACAGCGCGCACAGAT...TATTTGG 952
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
311 yAspProLeuThrGlyArgLeuGlyLysLysGluGluGluProPheLeuG 328
953 GACGCTACCAACAATCAGATTCGCTATTCGAAGAAGCGCGCACAAAG 1002
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
328 lPheAlaArgAspHisSerIleSerValLeuHisAsnProThrLysArgG 344
1003 CTGTTGCGGTGGTTCGCCCGCACCGCAAAATACTCATCAACGGCTAC 1052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
345 LeuPheSerPheLeuArgIleGlyPheAsnLysProThrPheThrLys 361
1053 AACCTCGCGCATTTCTGAAAAACAAACTCTCAACTTCAC.....A 1096
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 rTyrlLeuSerGlyPhePheLysLysLysArgThrTyrlThrAsnProAsp 378
1097 GAGCGCTACAGCGCGCGACCGCCCATGTCGCGATTGTACTTACAG 1146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
378 hrAsnLeuHisGlyLuhThrArgProIleIleAspThrAspIleTyrlAsp 394
1147 CGCGTATGCGCTTGATATTCCTGCCACCGCTGTTGGCGGATTTAAT 1196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
395 LysValMetProMetArgIleProValAlaProLeuIleLysAlaValI 411
1197 CGTCGCGCATACCGACCGCGGAGGATGTGGTGTGCTGAATTTGAGC 1246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
411 eThrLysAsnPheAspLeuAlaAsnGluLeuGlyPheLeuGluValLysG 428
1247 AAGAGACCTCGCTTGTGACGCTTCGTCGCCGCGGCAATTAACATAC 1296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428 LysGluAspPheAlaLeuProThrIleLeuAspProSerLysThrGluMet 444
1297 GCGCGCGCTGTTGCGCAAAAGTGTGCAAAACATTAAGAAAGAA 1338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
445 LeuThrIleValLysGluSerLeuIleGluTyrlAlaLysGlu 458
seq_name: p1r2:F71489

```

```

seq_documentation_block:
sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain [siml1a
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 02-Mar-2001
C:Accession: F71489
C:Stephens, R.S.; Kalman, S.;ammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A:Reference number: A71570; M01D:99008080
A:Accession: F71489
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <ARN>
A:Cross-references: GB:AE001334; GB:AE001273; NID:93329078; PIDN:AMC68238.1; PID:9332
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: nqrA
C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino
C:Keywords: oxidoreductase

```

```

alignment_scores:
  quality: 446.00      Length: 464
  Ratio: 1.517        Gaps: 14
  Percent Similarity: 63.362      Percent Identity: 28.233

```

alignment block:

US-09-303-518d-125 x F71489 ..

Align seg 1/1 to: F71489 from: 1 to: 465

```

4 ATTTAAATCAAAAAAGGTCTTAAACCTGCCACATCGCGGCGACAGCCGAGCA 53
  ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
3 ILevalValSerArgGlyLeuAspSerLeuLysGlyAlaProLys61 19
54 AGCCCTTACGAGCGC.....CCGCGCATTTACCGAAGTCGCGTTGC 94
  : : : : : : : : : : : : : : : : : : : : : : : : : :
19 uSerGlyPheCysGlyLysValAspProThrTyrtValSerValAspLeu. 35
95 TTGGGAGAAATATGCCGCTATGCCCCCTCGATGAAGAAAGTCAAGAGAGC 144
  : : : : : : : : : : : : : : : : : : : : : : : : : :
36 ....ArgProPheAlaProLeuProLeuGlyValLysValThrProGlu 50
145 GATGCCGTCACAAAAAGCCAAAGTGTCTTTGAGACCAAAAGAAATCCGGC 194
  ||| ||| : : : : : : : : : : : : : : : : : : : :
51 AspGlnValThrAlaGlySerProLeuAlaGlyLysLeuPheSerG1 67
195 CGTGGCTTTACTGCGCGCTTACGCAAAATCGCGCATTCACCGTG 244
  ||| : : : : : : : : : : : : : : : : : : : :
67 yValPheLeuThrSerProValAspGlyGluValGlnLeuArgArg 84
245 GCGAAAGCGCGTACTTCACTCAGTCAGTCGATGCGCGTTGAA...GCGAAC 291
  ||| : : : : : : : : : : : : : : : : : : : :
84 LysnLysArgAlaLeuLeuGlnValLysLysLysProGlyLle 100
292 GACCAATTCAGATTGGAACGCTACGACCTGACACGCGTGCAGAAACTTAAG 341
  : : : : : : : : : : : : : : : : : : : :
101 SerGlnThrLysPheSer.....TyrAspLeuGlnSerLeuThr 113
342 CCGCGAAGAGATGCGCGCCCAACGTATCAATCCGTTGTGACTGCGC 391
  : : : : : : : : : : : : : : : : : : : :
113 rGlnLysAspLeuLeuGlnValPheLysLysGlnGlyLeuPheAlaLeuP 130
392 TGGCACCCTGCGTTCAGCAAAATCTGCGCTGCGTCCGATGCGCGAG...CCG 438
  : : : : : : : : : : : : : : : : : : : :
130 heLysGlnAlaProPheAsp...IleProAlaLeuProThrGlnSerPro 145
439 TTCGCCATCTTCGTCATCGATGACACACCAATCCGCTG.....GCTGC 482
  : : : : : : : : : : : : : : : : : : : :
146 ArgAspValPheIleAsnLeuAlaAspAsnArgProPheThrProSerVa 162
483 CGACCCCTACGCTATTCATAAGAGACCGCGCAGAT.....T 520
  : : : : : : : : : : : : : : : : : : : :
162 IGIuLysHisLeuSerLeuPheSerSerLysGlnAspGlyTyrThrIleP 179
521 TCAGACGCGCGCTGTGTGATTTGAGCGGTTTGACGAGCAAAATCCAT 570
  ||| : : : : : : : : : : : : : : : : : : : :
179 heValValGlyValGlnAlaIleAlaLysLeuPheGlyLeuLysProHis 195
571 GTTTGAGAGCAGCTGCGCGACAGCTGCGTGAATGTCGCCAATCAT 620
  : : : : : : : : : : : : : : : : : : : :
196 IleIleSerThrAspArgLeuThrLeuProThrGlnAspLeuValSerI1 212
  : : : : : : : : : : : : : : : : : : : :
621 C...GAACACATGATTCGCGCGCGCGCGCATCTCGCGTTGAGTGCA 667
  ||| : : : : : : : : : : : : : : : : : : : :
212 eAlaHisLeuHisThrIleAspGlyProPheProSerGlySerProSerT 229
668 CGCAGATTCATTTCATCGAGCCGCTC...GGCGCGAATTTAAACCGTGAG 714
  ||| : : : : : : : : : : : : : : : : : : : :
229 hrHisIleHisIleAlaArgIleArgAsnGlnArgAspValAlaPhe 245
715 ACCATCAATTCATCAAGATTAATTAACATGCGCTGTTGTTGTTGCAACAG 764
  ||| : : : : : : : : : : : : : : : : : : : :
246 ThrIleSerPheGlnGluValLeuSerIleGlyHisLeuPheLeuLysG1 262
765 CCGTCTGAACACCGAGCGCGGTGATGCGCTAGCGCTTCTCAAGTC.... 810
  ||| : : : : : : : : : : : : : : : : : : : :
262 yheValLeuGlnGlyGlnGlnIleValAlaLeuAlaGlySerAlaLeuProp 279

```

```

811 ..AACAAACCGCGCTCTTTCGCTACCGTTTGGGTGCGAAAGATTCGCA 858
  : : : : : : : : : : : : : : : : : : : :
279 roSerGlnAlaGlyLysTyrLeuIleThrAlaLysGlyAlaSerPheSerAsp 295
859 ATTATGCGGGCGCAATTTGTTGACACAGACAGAAC...CGCGTATTTCCG 905
  : : : : : : : : : : : : : : : : : : : :
296 LeuLeuProLysAspIlePheSerSerAspGlnIleThrLeuIleSerG1 312
906 TTGGGTATTGAACGCGCGCATTTACACAGAGCGCGCAGCAT...TATTTGG 952
  : : : : : : : : : : : : : : : : : : : :
312 yAspProLeuThrGlyArgLeuLysLysGlnGlnAsnProCysLeuG 329
953 GACGCTACCAACATTCAGATTTCGTTATCGAAGAAAGCCGCGACCAAGAG 1002
  ||| : : : : : : : : : : : : : : : : : : : :
329 LysMetArgAspHisThrIleThrLeuLeuProAsnProLysThrArgGln 345
1003 CTGTGGCTGGGTGGTGGCGCGACCGCGCAAAATCTCCATCACCCTGAC 1052
  ||| : : : : : : : : : : : : : : : : : : : :
346 SerPheSerPheLeuArgLeuGlyTyrAsnLysLeuThrValThrArgTh 362
1053 AACCCCTGGCGCATTTCTGAAAAACAA.....CTCTTCAAGTTCAACA 1096
  ||| : : : : : : : : : : : : : : : : : : : :
362 rTyrLeuSerGlyPhePheLysArgLysArgValPheMetAspMetAspT 379
1097 CAGCCGTCAACGCGCGCGACCGCGCATGTGCGCATTTGTCTATACAG 1146
  ||| : : : : : : : : : : : : : : : : : : : :
379 hrAsnMetIleGlyGlnLysArgProIleIleAspAlaGlnIleLysGln 395
1147 CCGGTGATGCCCTTGGAATTCCTGCCACCCCTGCTTTTGGCGGATTTAT 1196
  ||| : : : : : : : : : : : : : : : : : : : :
396 ArgValSerAlaIleProValProValAlaLeuIleIleLysAlaLeuG1 412
1197 CGTGGCGCATTCACGACAGCGCGCAGCATTTGGTGTCTGGAATTTGACG 1246
  : : : : : : : : : : : : : : : : : : : :
412 uThrGlnAsnPheGlnGlnAlaCysArgLeuGlyLeuLeuValAlaP 429
1247 AAGAAGACCTCGCTTTGTGACAGCTGCTGCGCGCGCAATTAAGAAATAC 1296
  ||| : : : : : : : : : : : : : : : : : : : :
429 roGluAspPheAlaLeuProThrPheIleAspProSerLysThrGlnMet 445
1297 GGCCCGCTGTGGCGCAAAAGTCTGGAANCAATTGAGAGGAA 1338
  : : : : : : : : : : : : : : : : : : : :
446 PheSerIleValLysGlnSerLeuLeuArgTyrAlaLysGln 459

```

seq_name: p1r2:E82252

seq_documentation_block:

```

Info-related protein VC1015 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C/Accession: E82252
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
  chardson, D.; Ermolaeva, M.D.; Yamachyan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
  1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833
A/Accession: E82252
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-801 <HEID>
A/Cross-references: GB:AE004183; GB:AE003852; MID:g9655473; PIDN:AAF94176.1; GSPDB:GN
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A:Gene: VC1015
A:Map position: 1
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S]

```

alignment_scores:

```

Quality: 158.00 Length: 472
Ratio: 0.721 Gaps: 16
Percent Similarity: 46.398 Percent Identity: 20.127

```

alignment_block:
US-09-303-518D-125 x E82252

Align seg 1/1 to: E82252 from: 1 to: 801

```

52 CAAGCCGTTTACGAGCGCCGCGCATTCAGCAATCGCTGTCGCGCA 101
   ||| : : : : : : : : : : : : : : : : : : : : : : : :
59 GlnProIleIleAsnAlaSerIleProAsnGlnLeuValLeuProLeu 75
   : : : : : : : : : : : : : : : : : : : : : : : :
102 AGAATATGCGCGTATGCGCCCTCGATGAAGATCAAGAAAGCGATCCG 151
   : : : : : : : : : : : : : : : : : : : : : : : :
75 sglHisIleIleGlyLysAlaGlyLeuValLysValGlyAspArg 92
   : : : : : : : : : : : : : : : : : : : : : : : :
152 TCAAAAAGGCGCAGTGTCTTTGAGACAAAGAAATCCGGCGGTGTG 201
   ||| ||||| ||| : : : : : : : : : : : : : : : :
92 alleuLysGlyGlnProLeuThrGlnIleThrSerThrPheMetLeuPro 108
   : : : : : : : : : : : : : : : : : : : : : : : :
202 TTTACTGCGCGCGCTTACGCAAAATCGCGCATTCACCGTGGCGAAA 251
   : : : : : : : : : : : : : : : : : : : : : : : :
109 IleHisAlaProThrSerGlyValIleSerAlaIle.....GluPr 122
252 CGCGGCTACTTCAG.....TCAGTGTGATTCG 277
   ||| : : : : : : : : : : : : : : : : : : : : : :
122 oArgThrValAlaHisProSerGlyLeuSerGluLeuCysIleValLeu 139
   : : : : : : : : : : : : : : : : : : : : : : : :
278 CCGTTGAGGCAAGCAGACGAA.....ATGAGTTT 306
   : : : : : : : : : : : : : : : : : : : : : : : :
139 hrProAspGlnGlnGlnIleThrPheGlnLeuGlnProGlnProAspPhe 155
   : : : : : : : : : : : : : : : : : : : : : : : :
307 GAAGCGTACGACCTGAAGCGCTGGCAAACTTAAGCGCGAAGAGTGC 356
   : : : : : : : : : : : : : : : : : : : : : : : :
156 GlnGlnLeuThrProGlnThrLeuGlnLeuIleArgGlnIleGlyIle 172
   : : : : : : : : : : : : : : : : : : : : : : : :
357 CCGCAACTGATTCATCCGCTTGTGACTGCGCTGCGACCCGCTCCGT 406
   : : : : : : : : : : : : : : : : : : : : : : : :
172 eSerGlyMetIleGlyAlaGlyPheProThrAla..... 183
407 TCAGCAAAATCTCGCGTGCATGCGCGACGCGTTCGCCATCTTGTCAT 456
   : : : : : : : : : : : : : : : : : : : : : : : :
184 ..LysLysLeuGlnSerGlyLeuSerArgThrGlnIleLeuIleAsn 199
   : : : : : : : : : : : : : : : : : : : : : : : :
457 GCGATGACACCAATCCGCTGGTGCACGACCTTACGGTATTCATCAAGA 506
   ||| : : : : : : : : : : : : : : : : : : : : : :
200 AlaIleGlnCysGluProGlyIleThrAlaAspValLeuMetArgGln 216
   : : : : : : : : : : : : : : : : : : : : : : : :
507 AGCGCGGAGGATTCAAAGCGCGCTGTTGATTCAGCGCGTTTGACG 556
   : : : : : : : : : : : : : : : : : : : : : : : :
216 nTyrAlaHisGlnIleIleGlnGlyIleGlnIleValGlnHisIleLeu 233
557 AACGCAAAATCCATGTT..... 573
   : : : : : : : : : : : : : : : : : : : : : : : :
233 yspProLysLeuThrIleIleGlyIleGlnAspAsnLysProGlnAlaVal 249
574 .....TGTAAAGCAGCTGGCGCAGACGCTGCT..... 603
   : : : : : : : : : : : : : : : : : : : : : : : :
250 AlaIleLeuGlnGlnAlaIleAlaGlnAspLysProMetValIleArgVal 266
604 .....GAAATATCTG 613
266 eProThrLysTyrProSerGlyGlyLysGlnLeuIleLysIleLeu 283
614 CCAACATCGAACAACATGATTCGCGCGCGCATCTCTGCGGTTTGAGT 663
   : : : : : : : : : : : : : : : : : : : : : : : :
283 hrAsnLeuGlnValProLysGlyIleProAlaAspIleGlyLeu... 298
664 GGCACGCAATTCATTCATGAGCGCGCGCGCATTAACCAACCGTGTG 713
   : : : : : : : : : : : : : : : : : : : : : : : :
299 .....MetValGlnAsnValGlySer..... 305
714 GACCATCATTTATCAAGATGTAATTACCATGCGCGTTGTTGTCACAG 763
   : : : : : : : : : : : : : : : : : : : : : : : :
306 .....LeuGlnAlaIleIleAlaArgAlaIleValHisG 316

```

```

764 GCCTGTGAACACCGAGCGGTGATTGCCCTAGGTGTTCTCAAGTCAAC 813
   ||| : : : : : : : : : : : : : : : : : : : : : :
316 LysIleProLeuIleArgValIleThrLeuThrGlyAspPheArg 332
814 AACCGCGCGCTCTTGGTACCGTTTGGGTGCGAAAGTATCGCAATTCAC 863
   ||||| : : : : : : : : : : : : : : : : : : : : :
333 LysProArgAsnValThrPalLeuLeuGlyThrProValGlnAlaLeu 349
864 TCGCGCGCAATTGTTGACACAGACAC.....CGCGTATTCG 904
349 unsngIlePheGlyTyrLysAlaAspLysLysLeuProArgLeuIleMetG 366
905 GTTCGATTGAACGCGCGCATTCACAGAGCGCGCATTAATTGGA 954
366 LysIleProMetIleGlyPheThrLeuPro..... 375
955 CGCTACCAATTCAGATTCGTTATCGAAGAGCGCGCAAGAGCT 1004
   ||| : : : : : : : : : : : : : : : : : : : : :
376 .....HisAlaGlnValProIle.....ThrArgThrAl 385
1005 GTTCGCGCTGGTTCGCGCGCGACCGACAAATATCTCATCATCGCTACA 1054
   : : : : : : : : : : : : : : : : : : : : : : : :
385 asnCysIleLeuAlaPro.....ThrArgAsnG 395
1055 CCCTCGCGCATTTCTGAAGAACTCTTCAGTTCAACACACCGCTC 1104
   ||| : : : : : : : : : : : : : : : : : : : : :
395 LysLeu.....ThrSer 398
1105 AACGCGCGCGCGCGCGCATGTGTCGATGTTGACTTACGACGCGGTGAT 1154
   : : : : : : : : : : : : : : : : : : : : : : : :
399 SerAspAsnGlnMetAlaCysIleArgCysGlyGlnCysAlaGlnAlaC 415
1155 GCCTGTGATTCCTCCACCCCTGCTTGTGGCGCATTAATCGTCGCG 1204
   ||| : : : : : : : : : : : : : : : : : : : : :
415 sProValSerLeuLeuProGlnGlnLeuGlnIleThrPheAlaLysAlaGln 432
1205 ATACCGACGAGCGCGCGCATTTGGTTCGTTGATTCGACGAGCAAGAC 1254
   : : : : : : : : : : : : : : : : : : : : : : : :
432 LysPheAspLys.....CysGlnGlnLeuAsnLysPheLysAsp 443
1255 .....CTGCTTGTGACAGCTTGTGTCGCGCGCGCAATTCAGAA 1295
444 CysIleGlnCysGlyAlaCysAlaTyrValCysProSerGlnIle..... 458
1296 CGGCGCGCTGTGGCG 1311
459 ....ProLeuValGln 462

```

seq_name: p1r2:S39893

seq_documentation_block:

```

rnfC protein - Rhodobacter capsulatus
C:Species: Rhodobacter capsulatus
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 17-Mar-2000
C:Accession: S39893
R:Schmehl, M.; Jahn, A.; Meyer zu Vilsendorf, A.; Hennecke, S.; Schuppel
Mol. Gen. Genet. 241, 602-615, 1993
A:Title: Identification of a new class of nitrogen fixation genes in Rhodobacter caps
A:Reference number: S39892; M01D:9408454
A:Accession: S39893
A:Molecule type: DNA
A:Residues: 1-519 <SC>
A:Cross-references: EMBL:X72888; NTD:9435523; PIDN:CA51399.1; PID:9435525
C:Gene: rnfC
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S]
C:Keywords: iron-sulfur protein
P:374-438/Domain: ferredoxin 2[4Fe-4S] homology <FER>

```

alignment_scores: Quality: 150.00 Length: 454

Quality: 144.50 Length: 458
 Ratio: 0.698 Gaps: 18
 Percent Similarity: 45.197 Percent Identity: 21.616

Alignment_block:
 US-09-303-518D-125 x AE0692 ..

Align seg 1/1 to: AE0692 from: 1 to: 673

```

100 GAGAGATATGCGCGGTATGCGCCCTCGATGAAGTCAAGAGAGCGCATGC 149
    : : : : : : : : : : : : : : : : : : : : : : : : : :
48 LysGlnHisIleGlyAlaGlyIleuGlyValSerValGlyAspArg 64
150 CGTCAAAAAGAGCGAGTGTGTTGAAGCAAAAAGAGATCGCGCGGTGG 199
    ||| : : : : : : : : : : : : : : : : : : : : : : :
64 gvalLeuArgGlyGlnAlaLeuThrArgGlyArgGlyArgMetLeuProV 81
200 TGTTTACTGCGCGCGGTTCAGGCAAAATCGCGCGCAT 237
    ||| : : : : : : : : : : : : : : : : : : : : : : :
81 alHis...AlaProThrSerGlyThrValIleAlaIleAlaProHisSer 96
238 .....CACCGTGGCGAAAGCGCGTACTTCAGTCAGTCGATGGCGGT 281
    ||| : : : : : : : : : : : : : : : : : : : : : : :
97 ThrAlaHisProSerAlaLeuAlaGlyLeu...SerValIleIleAspAl 112
282 TGAAGCAAGCAGC..... 294
112 asproIleuAspArgTrpIleGlyArgGlyGlyTrpSerAspArg 129
295 .....GAAATCGAGTTTGAACGCTAGCGACCTGAAGCGGTGCA 333
    ||| : : : : : : : : : : : : : : : : : : : : : : :
129 IahIserArgGlyAlaLeuIleGlyIleuArgGlyIleGlyIleGlyAla 145
334 AACTTAAGCGGCGAAGAGTGGCGCGCAACCTGATCCATCGCGTTGTG 383
    : : : : : : : : : : : : : : : : : : : : : : :
146 GlyLeuGlyGly.....AlaGlyPheArg 153
384 GAGTGGCGTGGCGCGCGCGTTCAGCAAAATTCCTGCGCGGTGCGCG 433
    ||| : : : : : : : : : : : : : : : : : : : : : : :
153 oThrGlyAla.....LysLeuGlnGlyGlyAspArg 164
434 AGCGGTTCGCGCATCTTCATAGCGATGAGCAGCAATCGCGTGGTGGC 483
    : : : : : : : : : : : : : : : : : : : : : : :
164 yslIleThrThrLeuIleIleAsnAlaIleGlyGlyProGlyThrIle 180
484 GACCTACCGTCAATTATCAAGAGCGCGCGGATTCGAACGCGCGGT 533
    : : : : : : : : : : : : : : : : : : : : : : :
181 AlaAspAspArgLeuMetGlnAspCysAlaAlaGlnIleValGlyIle 197
534 GTTGTATGAGCGCGTTGACCGAA..... 558
197 eArgIleLeuAlaHisIleLeuGlnProArgGlyValLeuIleGlyIle 214
559 .....CGCAAAATCCATGTTTGAAGCGACGTCGCGCAGC 594
    : : : : : : : : : : : : : : : : : : : : : : :
214 IuAspAsnLysProGlnAlaIleSerMetLeuArgAlaValLeuAlaAsp 230
595 GTGCGCGTGGAAAATGCTGCGCAACATCGAAGATTCGCGCGCGCC 644
    ||| : : : : : : : : : : : : : : : : : : : : : : :
231 .....AlaHisAspIleSerLeuArgValIleProThrIle 242
645 GCATCTGCGCGGT.....TTGAGTGGCAGCG 670
    : : : : : : : : : : : : : : : : : : : : : : :
242 sTyfProSerGlyGlyAlaGlyGlnLeuThrGlnIleLeuThrGlyGly 259
671 ACATTCATTTTCATCGAGCGCGTGGCGCGAATAAACGCTGTGAGCATC 720
    : : : : : : : : : : : : : : : : : : : : : : :
259 IuVal.....ProHisGlyGlyArgSerSerAspIleGlyVal 271
721 AATTATCAAGATGTAATTAC.....ATTGCGCGTTGTGTTCCAAC 761
    ||| : : : : : : : : : : : : : : : : : : : : : : :
272 LeuMetGlnAsnValGlyThrAlaGlyAlaValLysArgAlaValAlas 288
  
```

```

762 AGCGCGTTCAGAACCGGCGCGGTATTCGCTAGCGTTCAGATCA 811
    ||| : : : : : : : : : : : : : : : : : : : : : : :
288 pGlyGluProIleThrGlyArgValAlaThrIleThrGlyAlaIle 305
812 ACNACCGCGCGCTTTCGCTACCGTTTGGTGGCGAAATATCGCAAT 861
    : : : : : : : : : : : : : : : : : : : : : : :
305 eArgProGlnAsnValIleThrAlaArgLeuGlyThrProValArgHisLeu 321
862 .....ACTGCGGCGCAATTTGGTACACAGACAAACGCGGTATTCGCG 905
    ||| : : : : : : : : : : : : : : : : : : : : : : :
322 LeuAsnAspAlaGlyPheCysProSerAlaAspGlnMetValIleMet 338
906 TTGCGTATTTGAACGCGCGCATTAACAGAGCGCGCAGATATTGGGAC 955
    ||| : : : : : : : : : : : : : : : : : : : : : : :
338 yGlyProLeuMetGlyPheThrLeuProThrLeuAspValProValAla 355
956 GCTACCGCAATCAGATT.....TTCGTTTCGAAAGAGCGCGCGCAC 996
    : : : : : : : : : : : : : : : : : : : : : : :
355 yslIleThrAsnCysLeuLeuAlaProSerValThrGlyMetGly..... 369
997 AAGAGCGTGTTCGCGTGGTGGCGCGCGCAGCGCAAAATCTCATTCAC 1046
    ||| : : : : : : : : : : : : : : : : : : : : : : :
370 .....AlaProGlnGlyGlyLysSerCysIleArg 379
1047 GCGTACAAACCTCGCCCATTTCTGAAAACAAACTTTCAGTTCAACA 1096
    : : : : : : : : : : : : : : : : : : : : : : :
379 gCysSerAlaCysAlaAsp..... 385
1097 CAGCGCGTCAACGCGCGCGCGCGCATGTTGCGCATTTGTTACTTACAG 1146
385 ..... 385
1147 CGCGTATGCGCTTCGATATCTGCGCACCGCTGTTTGGCGCATTTAAT 1196
    : : : : : : : : : : : : : : : : : : : : : : :
386 ...AlaCysProAlaAspLeuLeuProGlnGlnLeuThrTrpPheSer 401
1197 CGTGGCGATTCAGCAGCGCGCAGCGCA.....TTGCGTTCCTGG 1237
    : : : : : : : : : : : : : : : : : : : : : : :
401 sGlyGlnGlnHisAspCysAlaThrAlaHisIleAlaAspCysIleG 418
1238 AATTGAGCAAGAGAGACCTGCTTGGCAGCTTCGTCGCGCGCGCAAA 1287
    ||| : : : : : : : : : : : : : : : : : : : : : : :
418 Iu.....CysGlyAlaCysAlaThrValCysProSerAsn 429
1288 TACGATACGCGCGCGCTGTGCGC 1311
    ||| : : : : : : : : : : : : : : : : : : : : : : :
430 Ile.....ProLeuValGln 434
  
```

seq_name: p1r2:E72398

seq_documentation_block:

hypothetical protein TM0244 - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: E72398
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, J.R.; Garret, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; M01D:99287316
 A:Accession: E72398
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1451 <ARN>
 A:Cross-references: GB:AE001708; GB:AE000512; NID:94980740; PIDN:AAD5335.1; PTD:9498
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM0244

alignment_scores:
 Quality: 140.00 Length: 454

Ratio: 0.648 Gaps: 17
Percent Similarity: 47.577 Percent Identity: 20.264

alignment_block:

US-09-303-518D-125 x E72398

Align seg 1/1 to: E72398 from: 1 to: 451

```

49 GAGCAAGCCGTTTACGACGCGCCGCGCATTTACGAAGTCGCGTTCGTTGG 98
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
35 AspLysProIleGluArgLalProLeuProGlnLysValAlaPheValPheLe 51
   ::::: ::::: ::::: ::::: ::::: ::::: :::::
99 GCAAGAAATATGCGCGGTATGCGCCCTCGATGAAATCAAGAGAGCGCATG 148
   ::::: ::::: ::::: ::::: ::::: ::::: :::::
51 uSerSnhIsalaglYasnProIalalYasnProValYalSerProglYasn 68
   ::::: ::::: ::::: ::::: ::::: :::::
149 CGGTCAAAAAAGGCCAGTGTGTGTGAAGCAAAAAAGAAATCCGGCGCTG 198
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
68 LuValLysThrGlyGlnValIleGlyGlnProGlnGlyPheIleSerAla 84
   ::::: ::::: ::::: ::::: ::::: :::::
199 GTGTTTACTGCGCGCGCTTCAAGGCAAAATCGCCGCGATTCACCGTGCGCA 248
   ::::: ::::: ::::: ::::: ::::: :::::
85 TyrLeuHISerProValThrGlyArgValLeuIle.....Guileil 97
   ::::: ::::: ::::: ::::: :::::
249 AAAGCGCGTACT.....CAGTCAGTCGCGA 274
   ||||| ||||| ||||| ||||| ||||| |||||
98 .LysLysIleLeuHISProIleLeuGlyLysProIleGluAlaIleVal 114
   ::::: ::::: ::::: ::::: :::::
275 TTGCGGTGAAGGCAAGCAGCAAAATC.....GAGTTT 306
   ::::: ||||| ||||| ||||| ||||| |||||
114 leGlLargThrSerAspSplurPrValHISleGluThrGlyAspPhe 130
   ::::: ||||| ||||| ||||| ||||| |||||
307 GAAGCCTACGCACTGGAAGCGCTGGCAAACTTAAGCGCGAAGAGTGGC 356
   ||||| ||||| ||||| ||||| ||||| |||||
131 GluArgMetSerLysGlnLulLeu.....Guileil 142
   ::::: ||||| ||||| ||||| ||||| |||||
357 CCGCAACCTGATCCAAATCGGTTGTGAGACTGCGTGGCAACCGCTCGCT 406
   ::::: ||||| ||||| ||||| ||||| |||||
142 elYelYalaglYalLeValGlyLeuGlyAlaMetCpheProThrHisV 159
   ::::: ||||| ||||| ||||| ||||| |||||
407 TCAGCAAAATTCCTGCGCGTGCATGCGAGCGCTGCGCATTCCTGTCAT 456
   ::::: ||||| ||||| ||||| ||||| |||||
159 alYsLeuSerProProGlnLysLysValAspThrLeuIleValasn 175
   ::::: ||||| ||||| ||||| ||||| |||||
457 GCGATGACACCAATTCGCTGGCGTCCGACCCCTACGTCATTAACAAGA 506
   ::::: ||||| ||||| ||||| ||||| |||||
176 GlYalaglYasnProValLeuThrIleAspHisArgLeuMetLeuG 192
   ::::: ||||| ||||| ||||| ||||| |||||
507 AGCGCGGAGGATTTCAACGCGCGCTGTGTGATTCAGCGCTTGGACG 556
   ||||| ||||| ||||| ||||| ||||| |||||
192 uArgLaglYasnProValLeuThrIleAspHisArgLeuMetLeuG 207
   ::::: ||||| ||||| ||||| ||||| |||||
557 AACCAAAATTCATGTTTGAAGCA...CGTGGCGCAGACGTCGCTCT 603
   ::::: ||||| ||||| ||||| ||||| |||||
208 .....LeuGlyValGlnLysAlaValAlaGlyValGlnSerAsnLys 221
   ::::: ||||| ||||| ||||| ||||| |||||
604 GAAATCTGCTCAACATGAAACACATGATTCGGGGCGCCGATCTGCG 653
   ::::: ||||| ||||| ||||| ||||| |||||
222 MetAspLalYrHisasnLeuLysLysValPheLysGly...TyrProVa 237
   ::::: ||||| ||||| ||||| ||||| |||||
654 CGGTGTAGTGGCAGCAGCATTCATTCATCGAGCGGTGCGCGCAANA 703
   ::::: ||||| ||||| ||||| ||||| |||||
237 lAspValAlaLeuLeuArgThrLysTyr.....ProGlnGlyAlaGlnL 252
   ::::: ||||| ||||| ||||| ||||| |||||
704 AA...ACCGTGGGACCATC..... 720
   ::::: ||||| ||||| ||||| ||||| |||||
252 yselnLeuIleYrAlaIleThrGlyArgMetValProArgLysGlyLeu 268
   ::::: ||||| ||||| ||||| ||||| |||||
721 .....AATTCAGAGATGATTAATACAT 743
   ::::: ||||| ||||| ||||| ||||| |||||
269 PrometAspValGlyValValValGlnasnValGlyThrCysValAlaLa 285
   ::::: ||||| ||||| ||||| ||||| |||||
744 TGCGCGTTGTTTGCACAGCGCGTCTGAACACCGAGCGCGTGAATGGCC 793

```

```

285 lLysGlnAlaValValAspLysProLeuValGluArgLysMetThrV 302
   ::::: ||||| ||||| ||||| ||||| |||||
794 TAGGTGTTCTCAAGTCACAAACCGCGCTTGTGCGTACCGTTTGGT 843
   ::::: ||||| ||||| ||||| ||||| |||||
302 alSerGlyAspAlaValLysAsnGlnLysasnLeuIleValArgIleGly 318
   ::::: ||||| ||||| ||||| ||||| |||||
844 GCGAAATATCGCAATTACT.....GCGGCGAAATGGTGACACAGA 887
   ::::: ||||| ||||| ||||| ||||| |||||
319 ThrProValLysaspValIleAspTyrCysGlyGlyIleaspLysasnH 335
   ::::: ||||| ||||| ||||| ||||| |||||
888 CAACCGCGTATTCGCGTTCGATTCGAACGCG...GGATTACACAG 934
   ::::: ||||| ||||| ||||| ||||| |||||
335 rGluArgValIleLeuGlnGlyLysProMetGlyIleSerIleThr... 350
   ::::: ||||| ||||| ||||| ||||| |||||
935 GCGCGCACGATTAATTGGAGCTACCAATCAATTCGATTATCGAA 984
   ::::: ||||| ||||| ||||| ||||| |||||
351 .....AsnLeuAspIleProValMet 357
   ::::: ||||| ||||| ||||| ||||| |||||
985 GAAGCGCGCAGCAAGAGCTGTTGCGCTGGGTTGCGCGCAGCGAGCA 1034
   ::::: ||||| ||||| ||||| ||||| |||||
358 LysGlyThrSer..... 361
   ::::: ||||| ||||| ||||| ||||| |||||
1035 ATACTCCATCACGCGTACACCCCTCGCCATTCCTGAAAAACAACTCT 1084
   ::::: ||||| ||||| ||||| ||||| |||||
362 ....GlyIleThrAla.....PheLeuProLys... 370
   ::::: ||||| ||||| ||||| ||||| |||||
1085 TCAGTTCAACACAGCGCTCAACGCGCGCGCGCCGCTGTCGCAT 1134
   ::::: ||||| ||||| ||||| ||||| |||||
371 .....SerArgProGlnLysProCysIleArgCys 380
   ::::: ||||| ||||| ||||| ||||| |||||
1135 GGTACTTACGAGCGCGTATGCCCTTGATTCCTGATTCGCCACCGCTCTT 1184
   ::::: ||||| ||||| ||||| ||||| |||||
381 SerGlnCysValGlnValCysProMetAsnGlnProTyrLeuLeu 397
   ::::: ||||| ||||| ||||| ||||| |||||
1185 GCGCATTTAATCGTCGCGGATACGACGCGCGAGCGATGGCTTCT 1234
   ::::: ||||| ||||| ||||| ||||| |||||
397 rLeuLeuSerThrLysArgLysTyrAspGlnLalValGlnasnGlyLeuM 414
   ::::: ||||| ||||| ||||| ||||| |||||
1235 TGGAAATTGAGCAGAGACGCTGCTTGTGAGCTTGCTGCGCCGCGC 1284
   ::::: ||||| ||||| ||||| ||||| |||||
414 eAspCysIleGlu.....CysGlySerCysThrTyrThrCysProSer 428
   ::::: ||||| ||||| ||||| ||||| |||||
1285 AATATCAGATAC 1296
   ||||| ||||| ||||| ||||| |||||
429 LysIleGlnHIS 432

```

```

seq_name: p1r2:E64136
seq_documentation_block:
  rnfC protein homolog - Haemophilus influenzae (strain Rd KW20)
  C/Species: Haemophilus influenzae
  C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #extL_change 17-Mar-2000
  C/Accession: E64136
  R:Feischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
  ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
  ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
  Science 269, 496-512, 1995
  A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
  A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
  A:Reference number: A64000; M01D:95350650
  A:Accession: E64136
  A:Status: nucleic acid sequence not shown; translation not shown
  A:Molecule type: DNA
  A:Residues: 1-819 <TIGR>
  A:Cross-references: GB:U32841; GB:I42023; MID:q1574529; PIDN:AA02331.1; PID:q1574537
  C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S]
  F:371-435/Domain: ferredoxin 2[4Fe-4S] homology <FER>

```

alignment_scores:
Quality: 138.50 Length: 468
Ratio: 0.635 Gaps: 17

Percent Similarity: 46.581 Percent Identity: 19.658

Alignment block:

US-09-303-518d-125 x E64136 ..

Align seg 1/1 to: E64136 from: 1 to: 819

```

70 CCGGCATTCACGAGTGGCGTCTGGCCAGAAATATGCCGGATGCG 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
38 ProLeuGlyThrAspPheTyrIleProLeuLysGlnHisLeuGlyThrTh 54
120 CCCCTGATGAAAGTCAAGAGAGCGGATGCCGTAAGAAAGGCGCAATGTC 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
54 rGlysnLeuLeuIleLysGlnGlyAspTyrValLeuLysGlnGlnAlaL 71
170 TGTGTGAGACAAAAAGATCCGGCGTGTGTCTTACTGCGCGGCTTCA 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
71 euThrLysGlyAspGlyLeuArgMetLeuProValHisAlaProThrSer 87
220 GGCAAATGCGCGCGATTCACCGTGGCGAAAGCGGCTACTTCACTGCACT 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 GlyThrIle.....LysSerIleLysProLysTyrVal 97
270 CGTGATGCGCGTGAAGCGACAGCAAAATCGAG..... 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
97 lAathrHisProSerGlyLeuAspGlyProThrIleHisLeuGlnAla 114
304 .....TTGAACGCTACGACCTGAAAGCGCTGGCAAC 336
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114 spGlyLeuAspGlnTrpIleGlnArgAsnProIleAspSerPheSerThr 130
337 TTAAGCGCGAAGAGATGCGCGCAACCTGATCAATCCGGTTGTGAC 386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
131 LeuSerSerGlnGlnLeuIleHisLysIleTyrGlnAlaGlyLeu...Al 146
387 TGGCGTGGCGACCCGTCGCTC.....AGCAAAATTCCTGGCGGTG 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 aglyLeuGlyGlyAlaValPheProThrAlaAlaLysIleGlnSerAlaG 163
428 ATGCGAGCGCGTTCGCCATCTGTCATGATGCGATGAGCACCAATCGCTG 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
163 LnglnLysValLysLeuLeuIleLeuAsnGlyAlaGlnLysGlnProTyr 179
478 GCTGCGGACCTTACGCTCATTTATCAAGAAGCCGCGAGGATTTCAACG 527
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
180 lIethrCysAspAspArgLeuMetArgGlnArgAlaAspGlnIleLeuL 196
528 CGGCGTGTGGTATTTGAGCGGTTTGACGACGCAAAATCCATGTTGTA 577
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196 sGlyIleArgIleLeuArgTyrIleLeuHisProGln...LysValValI 212
578 AGGCAGCTGGCGACGACGTGCGCT..... 603
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212 lAalrIleGlnAspAsnLysProGlnAlaIleSerAlaIleArgAsnAla 228
604 ...GAAATGCTGCCAATCAATCAATGATTTGGCGCGCGCGATCC 650
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
229 LnglnGlyAlaAsnAspIleSerIleArgValIleProThrLysTyrPr 245
651 TGGCGTGTGAGTGGACGACGACCATTTGATTCATTCAGAG.....C 688
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
245 oSerGlyAlaThrLysGlnLeuIleTyrIleLeuThrArgIleGlnValP 262
689 CGGTGCGCGGAAATAAACCGTGTGACCATCAATATCAAGATTAAT 738
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
262 roSerGlyLysArgSerSerIleGlyValLeuMetGlnAsnVal... 277
739 ACATGCGCGCTTGTGTGCAAGCGCT.....CTGAC..... 774
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
278 .....GlyThrMetPheAlaIleLysArgAlaIleLysAsnAspGlnP 292
775 .....ACGAGCGGCTGATTCCTAGGTGTTCTCAAGTACAAACCGC 820
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

292 oLeuIleGlnArgValValThrLeuThrGlyAsnLysIleAlaGlnLysG 309
821 GCCCTTGCGTACCGTTTGGTGGCGAAATATCGCAAT.....ACT 864
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
309 LysAsnTyrTrpValArgLeuGlyThrProIleSerGlnIleLeuSerAsp 325
865 GCGGCGCAATTTGGTGTGACACAGAACCGCGCTATTTCGGTGTGATT 914
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
326 AlaGlyTyrGlnPheAspLysHisPheProIlePheAlaGlyLysProk 342
915 GAGCGCGCGATTAACACAGCGCGGACGATTTTGGACGCTACACCA 964
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 tMetGlyLeu..... 345
965 ATCAGATTTCCGTTATCGAAGAGCGCGCAAGAGCTGTGCTGG 1014
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
346 ..GluLeuProAsnLeuAsnAlaProValThrLysLeuValAsnCysLeu 361
1015 GTTGGCGCGGACCGGACCAATATCTCATCAGCGGTACACCTCGGCCA 1064
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 LeuAlaProAspTyrLeuGlnTyr..... 369
1065 TTTCCTGAAAAACAACCTTCAAGTTCAACACAGCGGTACAGCGCGG 1114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
370 .....AlaGlnProGlnAlaG 375
1115 ACCGCGCCATGTCGCGATTTGACTTACAGACCGCGGTGATCCCTTGAT 1164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
375 LnglnAlaCysIleArgCysSerSerCysSerAspAlaCysProValAsn 391
1165 ATCTGCGCCACCCGCTTGGCGGATTTAATGCTGCGCATACCGACAG 1214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
392 LeuMetProGlnGlnLeuTyrTrpPheAlaArgSerLysAspHisLysI 408
1215 CGCGCAGCGCATTTGGT.....TGCTTGGAATTTGACGAAAGAACACC 1255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
408 sSerGlnGlnLysAlaLeuLysAspCysIleGln.....C 420
1256 TCGCTTGTGACGCTGCTGCGCGCGCAATATCGAATACGCGCGCTG 1305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
420 YsglyIleCysAlaTyrValCysPro.....SerHisIleProLeu 433
1306 TTGCGC.....AAAGTGTGGAACCATTTAGAA 1334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
434 lIleGlnTyrPheArgGlnGlnLysAlaLysIleTrpGlnIleLysGly 450
1335 GGA 1338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
450 sGln 451
seq_name: p1r2:S69703
seq_documentation_block:
HKRI protein precursor - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YDR420w
C:Species: Saccharomyces cerevisiae
C>Date: 22-Aug-1996 #sequence_rev1stion 06-Sep-1996 #text_change 23-Mar-2001
C:Accession: S69703; A53382
R:Dietrich, F.S.
submitted to the EMBL data library, August 1995
A:Description: The sequence of S. cerevisiae lambda 3641 and cosmids 9461, 9831, and
A:Reference number: S69555
A:Accession: S69703
A:Molecule type: DNA
A:Residues: 1-1802 <DIE>
A:Cross-references: EMBL:033007; NID:9927685; PIDN:AAB6487.1; PID:9927691; MIPS:YDR4
R:Kasahara, S.; Yamada, H.; Mio, T.; Shitatori, Y.; Miyamoto, C.; Yabe, T.; Nakajima,
J. Bacteriol. 176, 1488-1499, 1994
A:Title: Cloning of the Saccharomyces cerevisiae gene whose overexpression overcomes
A:Reference number: A53382; MIMD:94156857
A:Molecule type: DNA
A:Residues: 1-581,'A',583-593,'A',595-1802 <KAS>

```

A:Cross-references: EMBL:S69101; NID:9545659; PIDN:AB30051.1; PID:9545660
 A:Experimental source: YNN25
 A:Note: sequence extracted from NCBI backbone (NCBIN:144410, NCBI:P:144411)
 C:Genetics:
 A:Gene: SGD:HKR1
 A:Cross-references: SGD:S0002828; MIPS:YDR420W
 A:Map position: 4R
 C:Keywords: calcium binding; glycoprotein; transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-1802/Product: HKR1 protein #status predicted <MAT>
 F:1483-1508/Domain: transmembrane #status predicted <TM>
 F:1645-1656/Domain: calcium binding #status predicted <CAL>

alignment_scores:
 Quality: 135.00 Length: 497
 Ratio: 0.560 Gaps: 22
 Percent Similarity: 48.491 Percent Identity: 22.736

alignment_block:
 US-09-303-518d-125/rev x 569703 ..

Align seg 1/1 to: 569703 from: 1 to: 1802

```

1337 TCCTTCATATGTTCCAGACACTTGGCAACAGCGCGGATTCGTA 1288
|||||.....|.....|.....|.....|.....|.....|.....|
390 SerAspThrValSerThrSerThrAspThrValProTyrSer 406
1287 TTGGCCGGGACAGACAGCTGCACAAAGCGAGCTCTTCGTCGTCAT 1238
|||.....|.....|.....|.....|.....|.....|.....|
406 rVal...HisSerThrPheValHis...AlaThrSerSerThrTyr 421
1237 CCAGCAACCCAGTCTGCGCGCTGCGTATCGCCGACGATTAATCG 1188
|||||.....|.....|.....|.....|.....|.....|.....|
421 leSerSerSerLeuTyrSerSerProSerLeuSerAlaSerValSer 437
1187 CGCAAAAGAGGTTGGCGAGATATCCAAAGGATCAGCGCTGTAACT 1138
|||||.....|.....|.....|.....|.....|.....|.....|
438 HisPheGlyValAlaProPheProSerAlaTyrIleSerPheSerVal 454
1137 ACCAATC.....GGCACATGCGCGCGCGCTGAGCGGTG 1097
|||||.....|.....|.....|.....|.....|.....|.....|
454 lProValAlaValSerSerThrTyrThrSerProSerAlaSerVal 471
1096 TGTGAACCTTAAGATTGTTTTCAGGAATGCGGAGGTGTACGC 1047
471 al.....|.....|.....|.....|.....|.....|.....|
1046 GTGATGAGTATTTGTCGGCTGCGGCGCAACCGCGAAGCTCTT 997
|||||.....|.....|.....|.....|.....|.....|.....|
472 .....ValProSerAlaTyrAlaSerSerProSerValProVa 484
996 GCTGGCGCTTCTTCGATACGGAATCTGATGTGAGGTGCCAAT 947
|||||.....|.....|.....|.....|.....|.....|.....|
484 lAlaValSerSerThrTyrThrSer.....|.....|.....| 493
946 AATGTGCGCGCTGTGTAACTCGCGCTCAATACGGAACCGGAATC 897
|||||.....|.....|.....|.....|.....|.....|.....|
494 ..ProSerAlaProAlaIleSerSerThrTyrThrSerProSer 509
896 ACGGCGTTGTGTGTCAAC.....AATTCGCGCGC 865
|||.....|.....|.....|.....|.....|.....|.....|
510 lAlaProValAlaValSerSerThrTyrThrSerProSerAlaPro 526
864 AGTAATTTGGATCTTC.....GCACCAAAAACGGTAC 830
|||||.....|.....|.....|.....|.....|.....|.....|
526 aAlaIleSerSerThrTyrThrSerProSerAlaProValAlaVal 542
829 GCAAGAGCGCGGTTTGTGACTTGAGACCA.....CCTAGGCAATC 786
|||.....|.....|.....|.....|.....|.....|.....|
543 .....SerSerThrTyrThrSerSerProSerAlaProAlaIle 556
785 ACGGCTCGGTGTTCAGACGCGCTGTTCGAACAAACGGCAATGTAA 736

```

```

557 SerSerThrTyrThrSerSerProSerAla.....ProValAlaVal 570
|||.....|.....|.....|.....|.....|.....|.....|
735 TACATCTTGAATATGATGCTCCACACGCTTTATTCGCGCGACGCGC 687
|||.....|.....|.....|.....|.....|.....|.....|
570 lSerSer.....ThrTyrThrSerSerProSerAlaPro 581
687 .....|.....|.....|.....|.....|.....|.....|
581 roValAlaIleSerSerThrTyrThrSerSerProSerValProValAla 597
686 ...TCGATGAAGAATGTCGTCCACTCAACACGCGACAGA..... 648
|||.....|.....|.....|.....|.....|.....|.....|
598 ValSerSerThrTyrThrSerSerProSerAlaProAlaIleSerSe 614
647 .....TCGGGCGCGCGCAATTCATGATGTTTCGATGTCGACATTT 605
|||.....|.....|.....|.....|.....|.....|.....|
614 rThrTyrThrSerSerProSerAlaProValAlaValSerSerThrTyr 631
604 CAGACGCGACGTCGCGCGAGCTGCTTACAAACATGATTTGGCTTG 555
|||||.....|.....|.....|.....|.....|.....|.....|
631 hrSerSerProSerAlaProAlaIleSerSerThrTyrThrSer 647
554 .....GTCAACGCGCTCAATTCACCAACAGCGCGCTT 523
|||||.....|.....|.....|.....|.....|.....|.....|
648 ProSerValProValAlaValSerSerThrTyrThrSerPro..... 662
522 GAATCTCTCGCGCTTCTTGAATATGACCGTAGGCTGGACCGACG 473
|||||.....|.....|.....|.....|.....|.....|.....|
663 .....SerAlaProAlaIleSerSerThrTyrThrSerSerProSer 678
472 GATTGCTGTCATGCGATGACGAAGATGCGCAACGCTGCGATCGACG 423
|||||.....|.....|.....|.....|.....|.....|.....|
678 alProValAlaValSerSerThrTyrThrSerSerProSerAlaPro 694
422 GCAGCAATTTTCTGAACGAGCGGTGCGACGCGCATCCACAAACG 373
|||.....|.....|.....|.....|.....|.....|.....|
695 Ala.....IleSerSerThrTyrThrSerSerProSerAlaPro 708
372 TTGCATCAGGTTGCGCGCATCTTCTGCGCGCTTAAGTTGCCACG 323
|||||.....|.....|.....|.....|.....|.....|.....|
708 lAlaValSerSerThrTyrThrSerSerPro.....SerAlaPro 721
322 CAGGTGCGTAGCGTTCAACATCGATTCGTCGTTGCTTCAACGCGCA 273
|||||.....|.....|.....|.....|.....|.....|.....|
721 roAlaIle...IleSerSerThrTyrThrSerSerProSerAlaPro 736
272 ACCACTGACTGAAGTACGCGCTTTCGCCACGATGAATCGCGCATTT 223
|||||.....|.....|.....|.....|.....|.....|.....|
737 AlaValSerSerThrTyrThrSerSerProSerAlaProAlaIle.. 752
222 GCCTGAAGCGCGCGCAGTAAACACACGCGCGGATTTTGTCTTCA 173
|||||.....|.....|.....|.....|.....|.....|.....|
753 .....SerSerThrTyrThrSerSerProSerAlaProValAlaVal 767
172 ACAGCACTTGCGCTTTTTCAGCGCATCGCTTCTGACTTCATTCAG 123
|||||.....|.....|.....|.....|.....|.....|.....|
767 erSerThrTyr.....ThrSerSerProSerAlaIleValValIle 780
122 GAGCGCATACGCGCATATCT.....TCGCCAAGCA 91
|||.....|.....|.....|.....|.....|.....|.....|
781 SerSerThrSerThrSerSerProTyrAspIleValTyrSerProSer 797
90 CGCGACTTCGTAATGCGCGG.....CCGTCGTAACGCGCTTCTCG 47
|||.....|.....|.....|.....|.....|.....|.....|
797 rPheAlaIleIleSerSerGlyTyrThrProSerProSerAlaSerVal 814
46 GTCTGCGCGCATGCGGAGTTTAGACCTTTTATTTTA 6
|||||.....|.....|.....|.....|.....|.....|.....|
814 lMetSerSerThrSerSerSerProTyrAspIleVal 827

```

seq_name: p1r2:T02345

seq_documentation_block:
 hypothetical protein KIA0324 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 05-Nov-1999
 C:Accession: T02345
 R:Rick, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
 re, J.; White, S.; Ueng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
 submitted to the EMBL Data Library, March 1998
 A:Description: Sequencing of human chromosome 16p13.3.
 A:Reference number: Z14664
 A:Accession: T02345
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1791 <RIC>
 A:Cross-references: EMBL:AC004493; NID:g2996648; PION:AAC08453.1; PID:g2996650
 C:Genetics:
 A:Map position: 16
 A:Introns: 1610/2; 1706/2
 A:Note: KIA0324

alignment_scores:
 Quality: 131.00 Length: 411
 Ratio: 0.697 Gaps: 14
 Percent Similarity: 45.742 Percent Identity: 25.061

alignment_block:
 US-09-303-518D-125 x T02345

Align seg 1/1 to: T02345 from: 1 to: 1791

```

12 CAAAAGAGTCTAAACCTGCCATCGCGGCGACGCGCAAGCCGTTT 61
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
776 ArgArgArgSerAlaSerSerProArgThrLysThrSerArgGly 792
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 ACGACGGCC.....GCCATTACCGAAGTCGCGTT.... 92
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
792 YArgSerProSerProLysProArgGlyLeuGlnArgSerArgSer 809
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
93 .....GCTGGCGAAGAAATATGCCGTATGCCGCC 122
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
809 eraArgArgGlyLysThrArgThrArgArgArgArgSerGlySer 825
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 CTCGATGAAGTCAAGAGCGATGCCCAAAAAGCCAAAGTCGT 172
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
826 SerGlnSerThrSerArgArgGlnArgSerArgSerArgGly 842
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 TTGAAGCAAAAAGATCCGGCGCTGTTTACTCGCGGCTTCAGGC 222
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
842 LThrArgArgArgArgGlyGlySerGlyTyrHis..... 853
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 AAAATCGCGCGGATTCACCGCGGAAAGGCGTACTTCAGTCAGTC 272
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
854 ..SerArgSerProAlaArgGlnArgSerArgThrSer..... 867
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
273 GATTGCGGTTGAAGCAAGCAAGAAATCGATTGAAGCTACGACCTG 322
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
868 .....ArgArgArgArgGlyArgSerArgThrProProth..... 879
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
323 AAGCGCTGGCAACTTAAGCGCGCAAGAGTGGCGCCCAACTGATCAA 372
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
880 .....SerArgLysSerArgSerArgThrSerProAlaProT 893
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
373 TCCGATTGTTGATCGGCTGGCGACCGCGCTTCAGCAAAATTCCTGC 422
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
893 rLysArgSerArgSerArgAlaSerProAlaThrHisArgArgSer 909
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
423 CGTGGATGCCGAGCGGTTCGCATCTTCGCAATGCGATGACCAATC 472
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
910 SerArgThrProLeuHisSerArgArg..... 919
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
473 CGTGGCTGGCGACCGCTACGTCTTTCAAAGACCGCGCGAGATTTC 522
   ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

920 .....SerArgSerArg..... 923
523 AAGCGCGCTGTGTGATTGAACCGTTTGACCGAAGCAAAATCCATGT 572
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
924 ..ThrSerProAlaSer..... 928
573 TTGTAGGACAGTGGCGGACGAGCTGCTGTAATAATGCTGCCAATCG 622
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
929 .....ArgArgArgSer..... ArgSerArg 935
623 AAACACATGAATTCGGCGCGCCGATCTCGCGTTTGAGTGCAACGAC 672
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
935 gThrSerValThrArgArgSerArgSerArgAlaSerProAlaSerA 952
673 ATTCATTTCATCGAGCC.....GTCGCGCGCGCAATRA 704
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
952 rArgArgSerArgSerArgThrProProValThrArgArgArgSerArg 968
705 AACCGTGTGACCATCATTAATTAAGATGTAATTAACATTCGCGCTTGT 754
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
969 SerArgThrProThrThrArgArgArgSerArgSerArgThrProPro 985
755 TTGCAACAGCGCGTCTGAACACCGAGCGCGATTCGCCATGAGTCT 804
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
985 LThrArgArgSerArgSerArgSerArgThrProProValThrArgArg 1002
805 CAAGTCAACAAACCGCGCTCTTCGCTACCGCTTTGG..... 841
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1002 eraArgSerArgThrSerProLeuThrArgArgArgSerArgSerArg 1018
842 .....GTGCAAGATATGCAAAATTAATTCGCGGCAATTCGTTGAC 885
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1019 SerProValThrArgArgArgSerArgSerArgThrSerProValThr 1035
886 GACACCGCGGTGATTCGCTTCGATTGAACGCGCGCATTAACAAG 935
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1035 gArgArgSerArgSerArgThrSerProValThrArgArgSerArg 1052
936 CGCGGACG.....ATTATTGGAGCGTACCAATCAGATTCGCTTA 979
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1052 eraArgThrProProAlaLeuArgArgSerArgSerArgThrProLeu 1068
980 .....TCGCAAGAGCGCGCGCACAAAGACGTTGCGTGGTGGCGCG 1023
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1069 LeuProArgLysArgSerArgSerArgSerProLeuAlaLeuArgArg 1085
1024 CAGCCGCAAAATCTCCATCAACGCGTACA..... 1054
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1085 gSerArgSerArgThrProArgThrAlaArgGlyLysArgSerLeuTh 1102
1055 .....CCCTCGGCGAATTCCTGAAAACAAACTTCGAAGTTCACAC 1099
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1102 rSerProProAlaLeuArgArgSerArgAlaSerGlySerSerSer 1118
1100 CCGTCAACGCGCGGACCGCGCCATGCGTGCAGA 1132
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1119 ArgSerArgSerAlaThrProProAlaThrArg 1129

```

seq_name: pir2:A40670

seq_documentation_block:
 nuclear envelope protein POM 121 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 05-Nov-1999
 C:Accession: A40670
 R:Hallberg, E.; Wozniak, R.W.; Blobel, G.
 J. Cell Biol. 122, 513-521, 1993
 A:Title: An integral membrane protein of the pore membrane domain of the nuclear envelope
 A:Reference number: A40670; MUID:93328754
 A:Accession: A40670
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1199 <HAL>

A:Residues: 1-774 <STO>
 A:Cross-references: GB:AE004770; GB:AE004091; NID:g9949633; PIDN:AG06879.1; GSPDB:GN001
 A:Experimental source: strain PMO1
 C:Genetics:
 A:Gene: PA3491

Alignment_scores:
 Quality: 129.00 Length: 488
 Ratio: 0.578 Gaps: 23
 Percent Similarity: 45.697 Percent Identity: 21.516

Alignment_block:
 US-09-303-518D-125 x D83208 ..

Align seg 1/1 to: D83208 from: 1 to: 774

```

28 CTGCCCCATCGGGGCGAGCCGAGCAACCGCTTACGACGGCCGCGCCAT 77
   |||||
23 LeuProIleGlnGlnAlaProLeuAlaGlnArgTyrIleValPro.... 37
   |||||
78 TACCGAATCGCGTGTGCTGGCGGAATAT...GCCGTAATGCGCCCT 124
   |||||
38 .....LeuGlnHisIleGlyAlaProAlaArgProC 49
   |||||
125 CGATGAAGTCAAGAGAGCGATGCGTCAAAAAAGCCAAAGTGTGTT 174
   |||||
49 yValIleVal....GlyAlaValIleuLysGlyGlnThrIleAla 63
   |||||
175 GAAGCAAAAAAGATCCGCGGTGTGTACTCGCCGCGCTTCAGCAA 224
   |||||
64 LeuProAspGlyThrValSerAlaIleuHisAlaProThrSerGlyTh 80
   |||||
225 AATCGCCGCGATT.....CACCGTGGCGAAAAAGCGCGTAC 259
   |||||
80 yValValAlaIleGlyAlaHisProTyrProHisIleSerGlyLeuProA 97
   |||||
260 TTCAGTCAATCGTGAATGCGCTTGAAGCAAGCAAGAA..... 297
   |||||
97 LeuProAlaIleValIleAlaSerAspGlyLeuGlnArgTyrThrGluLeu 113
   |||||
298 .....ATCGAGTTTGAACGCTACGACCTGAAGCGCTGGCAAACTT 338
   |||||
114 HisProCysProAspPheArgIleGlnSerProLeuAlaLeu..... 128
   |||||
339 AAGCGCGAGAAAGTCCGCC.....AACCTGATCCAAATCCG 376
   |||||
129 .....GluArgIleArgAlaAlaGlyIleGlyLeuGlyGlyAlaG 143
   |||||
377 GTTGTGGACTGCG.....CTGGCACCGCTGCTCAGCAAAATTCCT 420
   |||||
143 LysPheProThrAlaAlaLysLeuAlaAlaArgProAlaGlnLysIle... 158
   |||||
421 GCCGTGATGCCGAGCGCTTCCGATCTTGCATTCGATGCGATGACACCA 470
   |||||
159 .....HisThrLeuValIleAsnGlyAlaGlnCysGln 169
   |||||
471 TCCGCTGGCTGGCGACCTACGCTCATTTCAAGAAGCGCGGAGGATT 520
   |||||
169 uProTyrIleSerAlaAspAspLeuMetArgGlnArgAlaThrGlnV 186
   |||||
521 TCAACGCGCGCTGTTGTTATGAGCGT..... 549
   |||||
186 alleuGlyGlyIleAspIleLeuValGlnIleLeuGlyProGlnGlnVal 202
   |||||
550 .....TTGACCGAAGCAAAATCATGTTTGAAGCGAGTGGCG 590
   |||||
203 LeuValIleGlyLeuAspAspLysProGlnAlaIleAlaIleLeuGlyAl 219
   |||||
591 A.....GAGTGGCGCTCAAAATGATGCGCAATCGAAGCAAG 631
   |||||
219 AlaLeuGlyLeuArgProTyrArgIleValAlaLeuProThrArgTyrP 236

```

```

632 AATTCGGCGCGCCGCAT..... 648
   |||||
236 roSerGlyGlyGlnArgGlnLeuIleGlnLeuThrGlyArgGlnVal 252
   |||||
649 CTTGCC...GGTTGAGTGGCAGCAGCATTCATTTCATGACGCGGTCGG 695
   |||||
253 ProAlaAspGlyLeuProAlaAspIleGlyIleLeuGlyGlnAsnValGln 269
   |||||
696 CGCGAATAAACCGTGTGGACCATTCATTTCAAGTGTATTTACCATTTG 745
   |||||
269 yThrLeuAlaIleVal.....HisAspAlaValIleG 281
   |||||
746 GCCGTTGTTTGCACAGCGCGTGTGACACGCGCGGTGATTCGCTTA 795
   |||||
281 LysArgProLeuIleSer.....ArgIleThrThrLeu 291
   |||||
796 GGTGTTTCTCAAGTCAACAAACGCGCTTCTGCTACCGCTTTGGTGC 845
   |||||
292 AlaGlyGlyAlaLeuGlnArgProMetAsnValGlnAlaLeuIleGlyTh 308
   |||||
846 GAAGTATCGCAATTTACTCG.....GGCGAATTCG 877
   |||||
308 rProValHisGlyLeuLeuAlaPheAlaGlyLeuAlaGlnGlyArgLeu 324
   |||||
325 .....GluArgValLeuMetGlyGly..... 331
   |||||
928 ACACAGCGCGCGCAGATTTATTTGGACGCTACACCAATCAGATTTCCG 977
   |||||
331 ..... 331
   |||||
978 TATCGAAGAGCGCGCAGCAAGAGCTTTGCGCTGGTTGCGCGCAGC 1027
   |||||
332 .....PrometMetGlyPheAlaLeuProAsp. 340
   |||||
1028 CGGACAAATTCATCCATCGCGGTACACCCCTGGCCATTTCTGTAAC 1077
   |||||
341 .....LeuSerVal 343
   |||||
1078 AATCTTTCATAGTCAACAGACCGCTCAACGCGCGCGCGCCAGCATGT 1127
   |||||
344 ProLeuIleLysThrCysAsnGlyLeuLeuAlaGlyAspAlaThrGlnLe 360
   |||||
1128 GCG.....ATTGTACTTAAGAC 1147
   |||||
360 uProGlnProValProAlaMetProCysIleArgCysGlyAspCysAlaG 377
   |||||
1148 GCGTATGCCCTTGGATATCCGCGCCACCGCTTTCGCGGATTTATC 1197
   |||||
377 LysValCysProValSerLeuLeuProGln...GlnLeuHisPhePheAla 392
   |||||
1198 GTGCGGATTAACGACAGCGCGAGCATTTGGGTTGCTTGAATTTGAGCA 1247
   |||||
393 LeuGlyAsp...GlnHisGlnIleLeuLeuAlaHisAsnLeuPheAspCys 408
   |||||
1248 AGAAGACCTGCGCTTGTGCAAGCTTGTGCGCGGCGCAAAATACGATACG 1297
   |||||
408 sIleGlnCysGlyAlaCysAlaTyrValCysProSerSerIle..... 422
   |||||
1298 GCCGCGCTTGGCG 1311
   |||||
423 .....ProLeuValGln 426
   |||||

```

seq_name: p1r2:A84943

seq_documentation_block:
 membrane protein [imported] - Buchnera sp. (strain Aps)
 C:Species: Buchnera sp.
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: A84943
 R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
 A:Reference number: A84943; MUID:20445173
 A:Accession: A84943
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-473 <570>
 A:Cross-references: GB:AP000398; GSPDB:GN00144
 A:Experimental source: strain APS
 C:Genetics:
 A:Gene: rnc

alignment_scores:

Quality: 126.00 Length: 449
 Ratio: 0.575 Gaps: 22
 Percent Similarity: 48.775 Percent Identity: 20.045

alignment_block:

US-09-303-518D-125 x A84943 ..

Align seg 1/1 to: A84943 from: 1 to: 473

```

100 GAAGAATATGCCGATATGCCCTCGATCAAGTCAAGAGGCGATGC 149
    ::::||||| ::::|||||
13  AspAspTyrHisAsnValLys.....LeuArgValLysIleAsnGluLys 27
150 CGTCAAAAAGGCCAA...GTCTGTTTGAAGACAAAAGATCCGGCG 186
    ||| :||| :||| :||| :||| :|||
27  ValLeuArgGlyInProLeuIlePheSerAspAspPheAsn.....Y 42
197 TGTGTTTACTGCGCGCTCAGCAAAATCCGCCGATTCACCGTGGC 246
    || :||| :||| :||| :||| :|||
42  alProValHisAlaProThrSerLeuIleGluAsnIleCysPheAsn 58
247 GAAAAGCGGCTACTTCAGTCAGTCGTGATTCGCCGTTGAGGC..... 288
    :::: :::: :::: :::: :::: ::::
59  SerAspSerIleLysLysAsnIleLysIleValIleSerProAspTyrLe 75
289 .....AACGACGAATCGAGTTTGAACGCTACGAC 319
    ||| :||| :||| :||| :||| :|||
75  uAspGlnTTrpIleArgLeuAsnProIleLysAspTyrLysLysTyrAla 92
320 CTGAAGCGGCTGGCAACCTTAAGCGCGAAGAGTGGCGCCGACACCTGATC 369
    ||||| ||| :||| :||| :||| :|||
92  roGluLysLeuIleLysIle.....IleHis 100
370 CAATCCGGTGTGGACTGGCTGGCCACCCGCTCCGTTCC.....AG 410
    ||||| ||| :||| :||| :||| :|||
101 GlnSerGlyVal...ValGlyLeuGlyGlyGlnPheProSerSerLys 116
411 CAAATTCCTGCCGTCGATGCGCGCTGTCGATCTCGTCAATAGCGA 460
    :||| :||| :||| :||| :||| :|||
116  slYsIleIlePheSerIleAsnArgAlaHisThrLeuIleValAsnAlaV 133
461 TGGACACCAATCCGCTGGCTGGCCGACCTACGCGCTTATTAAGAAGCC 510
    :||| :||| :||| :||| :||| :|||
133  alGlnSerGlnProTyrIleThrSerAspAsnLysLeuIleTyrAsnHis 149
511 GCCGAGGATTTCAACGCGCGCTGTGATTCAGCGCGTTTACCGGACG 560
    :||| :||| :||| :||| :||| :|||
150  IleSerGlnIleLeuIleGlyLysIleIleCysTTrpIleThrLysIle 166
561 CAAATTCATGTTTGTAGGACGTCGGCGACAGCGTCCGCTCT..... 603
    ||| :||| :||| :||| :||| :|||
166  elYsThrValLeuIleAlaIleGlnGluAspAsnIleGlnSerIleSerL 183
604 .....GAAATGCTGCCACATGGAACACATGAAATTC 636
    :||| :||| :||| :||| :||| :|||
183  ysIleGlnHisLeuIleLysAsnLysSerLeuPheLysIleCysIleIle 199
637 GCGGCGCCGATCTGCGCGT.....TTGAG 662
    :||| :||| :||| :||| :||| :|||
200  LysLysLysTyrProAlaGlySerSerLysValLeuValLysSerLeuTh 216
  
```

```

663  TGGCAGCCACATT.....CATTCATCAGAGCCGTCGCGCA 700
    ||| :||| :||| :||| :||| :|||
216  rGlyLysGluValProHisGlyLysHisSerIleAsp...IleGly.... 230
701  ATAAACCGTGTGACCATCAATTATCAAGATTAATTACATTGCGCGT 750
    :||| :||| :||| :||| :||| :|||
231  .....TyrLeuIlePheAsnValAlaThrIlePheSerIleLysArg 244
751  TTGTTTGCACAGCGCGCTGCAACACGAGCGCGCTGATTCGCCCTAGTGG 800
    :||| :||| :||| :||| :||| :|||
245  AlaIleIleAsnGlyLysProLeuThrGluArgValValThrLeu.... 259
801  TTCTCAAGTCAACAACACCGCGCTCTGCGTACCGCTTTGGTGGCGAAG 850
    :||| :||| :||| :||| :||| :|||
260  .....MetSerAspLysAsnLeuLeuSerGlyAsnPheThrValArgI 274
851  TATCGCAA.....ATTACTGCGCGCAATTGGTTGACACA 885
    :||| :||| :||| :||| :||| :|||
274  IeGlyThrProIleLysTyrPheLeuThrSerAsnLysLeuLysGlnSer 290
886  GACACACCGGTGATTTCCGGTTCGTTATTAACGCGCGCATTTACACAAAG 935
    :||| :||| :||| :||| :||| :|||
291  .....PheIleAlaSerValTyrLeuGlyLys..... 299
936  CGGCGACGATTAATTGGGACGCGTACCACATGAGATTCGTTATACAG 985
    :||| :||| :||| :||| :||| :|||
300  .....ProPheMetGlyLysLysIleAsnAsnLeuAsn..... 310
986  AAGCGCGCAGCAAAAGAGCTTCGCGCTGGTTCGCCCGACCGGACAA 1035
    :||| :||| :||| :||| :||| :|||
310  ..... 310
1036  TACTCATTCACGCGGTACACCTCGGCCATTTCTGAAAACAACTCTT 1085
    :||| :||| :||| :||| :||| :|||
311  HisSerIleLeuLysLysThrAsnSerIlePheIleThrHisLysGln 327
1086  CAAATTCAACACAGCGCTCAACGCGCGCGCGCGCCATGGTGGCGGANTG 1135
    ||| :||| :||| :||| :||| :|||
327  uLys...AsnGlnSerIleSer.....GlnLysThrLysIleArgCysG 341
1136  GTACTTACGAGCGCGTATGCCCTTGATATCTGCCACACCTGCT... 1182
    :||| :||| :||| :||| :||| :|||
341  LysTyrCysSerTyrValCysProValAsnLeuLeuProGlnIleGln 357
1183  .....TGGCGGATTTAATCGCGGATTCGACGACGCGCGCGGCGAT 1226
    :||| :||| :||| :||| :||| :|||
358  TrpTyrIleLysAsnLysAsnHisValGlnThrLysLysHisTyrValLe 374
1227  GCGTTGCTGGAATTTGACGAAAGACCTTCGTTGACGCTTCGCT 1276
    :||| :||| :||| :||| :||| :|||
374  uAspCysIleGlu.....CysLysAlaCysGlnLysValC 386
1277  GCCCGGGCAATACGAAATACGCGCGCGCTGTTGGCAAAAGCTGGAA 1323
    ||||| :||| :||| :||| :||| :|||
386  ysPro.....SerTyrIleProLeuValLysTyrPheIleGln 398
seq_name: p1r2:T39903
seq_documentation_block:
serine-rich protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Nov-2000
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z21889
A:Accession: T39903
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-534 <LYN>
A:Cross-references: EMBL:AL033543; PIDN:CAA22127.1; GSPDB:GN00067; SPDB:SPBC215.13
A:Experimental source: strain 972h; cosmid c215
  
```

C:Genetics:
A:Gene: SPDB:SPBC215.13
A:Map position: 2
C:Superfamily: pig submaxillary mucin

Alignment_scores:
Quality: 126.00 Length: 415
Ratio: 0.603 Gaps: 15
Percent Similarity: 50.361 Percent Identity: 20.241

Alignment_block:
US-09-303-518D-125/rev x T39903

Align seg 1/1 to: T39903 from: 1 to: 534

```

1280 GGGCAGACGAAAGCTGCACAAAGCGAGCTCTTCTGTCATTCACAGCA 1231
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 GlyGlyThrProThrTyrGlyTyrAlaThrProThrSerSerSerG1 83
1230 ACCCAATGCCGTGGCGCTGTGCGTATGCCGACATTAATCCGCGCAAAA 1181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 uProSer1LeuPheSerGluSerAlaThrProSerGluThrAsnSerTyrS 100
1180 GCAGGGTGGCAGCATATCAAGGCGATCAGCGCTCGTAGTACCAATC 1131
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 erSer.....ProVal 103
1130 GGCACCAATGGCGGCGTCCGCCCTTGACGCGCTGTGTAATGAGAG 1081
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
104 SerSerTyr.....SerAspProAlaThrSerGlnLeuProSerSerTh 118
1080 TTTGTTTTCAGAAATGCGCGAGGCTGTACGCTGATGAGATTTGT 1031
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118 SerPhePheSerPro.....ThrSerSerGluTyrThrp 130
1030 CCGCGTGGCGGCGACACCGACGAGCTCTTCTGCGGCTCTTCG 981
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
130 ro.....SerSerThrGluSerSerSerLeuLeuAspProSerSer 143
980 ATACGGAATCTGATGTGTAGCGTCCCAATATGCGCGCCCTTG 931
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144 ValSerSer..... 146
930 TGTAAATCGCGCGCTCAATACCGAACCAGAAATCAGCGGTGTCTGT 881
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147 AlaLeuProSerSerThrSerValGlnVal.....SerIles 160
880 CAACCAATTCGCGCGCAGATTTTGCATTTCTTGACCAACCAACGGTA 831
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
160 erSerSerSer..... 163
830 CCGAAGAGCGCGGCTTGTGACTTGAGAACACCTAGGCAATCAGCG 781
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
164 .....LeuSerSerSerAspProLeuThrSerSerThrp 175
780 CTCGGTGTACAGCGCCTGTTCGAAACAGCGCAATGTAATTAAT 731
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175 eSerSerLeuSerSerSerThrSerSerSerGlnProSerVal..... 189
730 CTTGATTAATTAATGTCACACAGGTTTATTGCGCGCGCGCGTCGATG 681
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 .....SerSerThrSerSerSer 195
680 AATGATATGCGTGCACACTCAACCGGAGATGCGGCGCGCGCAATTC 631
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196 .....ThrPheSerSerAlaAlaProThrSerTh 205
630 ATGTGTTTCGATGTGGCAGCATTTCA.....GACGGCAGCTGTGGC 587
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
205 rSerSerSerTyrLeuSerSerSerSerValValSerSerSerSerSer 222
586 GACGTCCCTTACAACATGATTTTGGCTTGGTCAACCGCTCAATACC 537

```

```

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
222 roSerSerSerSerSerThrLeuThrSerSerSerLeuSerThrSer 238
536 AACAGGCCCGCTTGAATCCCTCGGGGCTTCTTGATATACCGTAGG 487
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 SerLeuProSerThrSerSerSerSer..... 248
486 GTGCGCAGCGCAGGATTTGTCATCCATTTGACGAGATGCGAAGC 437
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
249 .SerThrSerSerSerLeuSerSerSerSerSerSerThrIleAsnS 265
436 GCTGGCATCGACGCGAGAAATTTGCTGACGAGACGGGTGCGCAGCA 387
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265 erSerSerSerSerSerSerIleLeuSerSer.....SerSerSer 279
386 GTCCACAAACCGGATTCGATCAGTTGCGCGGACCTTCTTCGCGCTTA 337
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
280 SerSerSerPro.....ThrSerThrSerThrIleSe 291
336 GTTTCGACGCGCTTCAGTGCATGCGTTCACAACTGATTCGCTGTC 287
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 rSerSerSerSerSerSerSerSerProThrSerThrSerThrIleS 308
286 CTTCAACGCGCAATCAGACTGACTGAAATGACGCGCTTTCGCCACGCT 237
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
308 erSerSerSerSer.....SerSerPheSerSerThrLeu 321
236 ATGCGCGCGATTTGCTGACGCGCGCGCAGTAAACACACGCGCGGAT 187
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
322 SerSerSerSerSerSerSerSerSerPheSerSerSerPro..... 336
186 CTTTTCCTTCAACAGCACTTGCCCTTTTTCGACGCGCATCCCTTC 137
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
337 ....ThrSerSerSerThrIleSerSerSerSerSerSerSerS 352
136 TGACTTCATCAGAGCGCGCATACCGCAATTCCTGCGCAACGAGCG 87
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
352 erSerSerPheSerSerThrThrSerSerSerSerSerSerSerPhe 368
86 ACTTCGGTAATGCGCGCGCGCTGTAACGCGCTTGTCTCGCGTGTG 42
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
369 SerSerThrValSerSerSerSerSerThrSerSerSerThrLeu 383

```

seq_name: p1r2:S01359

seq_documentation_block:

salivary glue protein sgs-3 precursor - fruit fly (*Drosophila erecta*)
C:Species: *Drosophila erecta*
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Feb-1997
R:Martin, C.H.; Mayeda, C.A.; Meyerowitz, E.M.
J. Mol. Biol. 201, 273-287, 1988
A:title: Evolution and expression of the Sgs-3 glue gene of *Drosophila*.
A:Reference number: S01358; MUID:88332966
A:Accession: S01359
A:status: not compared with conceptual translation
A:molecule type: DNA
A:Residues: 1-328 <MAR>
C:Genetics:
A:Gene: Sgs-3
A:Cross-references: FLYBase:FBgn0012268
C:Superfamily: salivary glue protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-328/Product: salivary glue protein sgs-3 #status predicted <MAT>

Alignment_scores:

Quality: 125.00 Length: 267
Ratio: 1.000 Gaps: 12
Percent Similarity: 46.816 Percent Identity: 25.468

Alignment_block:

US-09-303-518D-125 x S01359

Align seg 1/1 to: S01359 from: 1 to: 328

```

29 TCCCATCGCGGCGAGCGAGCAAGCGTTACGAGCGCCG... 73
   |||||:|||||:|||||:|||||:|||||:|||||:
28 CysProLysArgThr...ThrProLysProCysThrThrAlaArgProth 43
74 .....CCATTACCGAAGTCGCGCTGCTGGCGAATAATGCCGAT 116
   |||||:|||||:|||||:|||||:|||||:|||||:
43 rCysAlaProValThrThrThrCysArgProProThrThrThrArgc 60
117 GCGCCCTCGATGAAGATCAAGAGGCGATCCGTCAAAAAGCCAG 166
   |||||:|||||:|||||:|||||:|||||:|||||:
60 ySPioProProThrThrThrArgCysProProProThrArgProAlaGlu 76
167 TGCCTTTGAAGACAAAAAATCCGCGCTGCTTACTGCGCCG... 214
   |||||:|||||:|||||:|||||:|||||:|||||:
77 CysThrAlaThrThrLysArgProThrAlaArgProThrThrAlaGlyth 93
215 ....CTTACGCAAAATCGCGCGATTCACCGCGAAGCGCGTACT 260
   |||||:|||||:|||||:|||||:|||||:|||||:
93 rThrValArgAlaThrThrLysArgAlaThrThrArgThrThrLysA 110
261 TCAGTCAGTCG...TGATTGCGCTTGAAGCAAGCAAGCAAA...TCAGT 304
   |||||:|||||:|||||:|||||:|||||:|||||:
110 rGAlaThrThrArgArgThrThrValAlaArgAlaThrThrLysArgAla 126
305 TTGAACGCTACGACCTGAAAGCGCTGCGCAACT...TAAGCGCGAGAA 351
   |||||:|||||:|||||:|||||:|||||:|||||:
127 ThrArgArgThrThrThrLysArgAlaProThrArgArgThrThrLys 143
352 GTGCGCGCAACCTGATCCATCCGCTT...TGCGACTGCGTCCG 395
   |||||:|||||:|||||:|||||:|||||:|||||:
143 sArgAlaThrThrArgArgAsnProThrArgArgThrThrThrArgArg 160
396 CACCGTCGCTTCAAGAAATTCCTGCGCGATGCGAGCGCTGCCCA 445
   |||||:|||||:|||||:|||||:|||||:|||||:
160 lAProThrLysArgAlaThrThrLysArgAlaThrThrArgArgAsnPro 176
446 TCT...TCGCAATGCGATGAGCAACCAATCCG... 475
   |||||:|||||:|||||:|||||:|||||:|||||:
177 ThrLysArgLysThrThrArgArgThrThrValArgAlaThrLysThr 193
476 .....TGCGCGCGACCGCTACGCTCATATCAAGA 506
   |||||:|||||:|||||:|||||:|||||:|||||:
193 rLysArgAlaThrThrLysArgAlaProThrLysArgAlaThrThrLysA 210
507 AGCCGCG...AGGATTTCAACCGCGCTGTGTGATGAGCGGT 550
   |||||:|||||:|||||:|||||:|||||:|||||:
210 rGAlaProThrLysArgValThrThrLysArgAla... 221
551 TGACCGAAGCAAAATCCATGTTGTAAGCGAGCGCGAGACGTCGCG 600
   |||||:|||||:|||||:|||||:|||||:|||||:
222 ...ProThrLysArgAlaThrThrLysArgAlaProThrLysArgAla 237
601 TCGAAATAGCTGCCAACATCGAACAACATGATGCGCGCGCGATCC 650
   |||||:|||||:|||||:|||||:|||||:|||||:
237 rThrLysArgAlaProThrLysArgAlaThrThrLysArgAlaProThr 254
651 TGCGGCTTGTAGTGCGACG... 670
   |||||:|||||:|||||:|||||:|||||:|||||:
254 ySAArgAlaThrThrLysArgAlaProThrLysArgAlaThrThrLysArg 270
671 ...ACATTCATTCATCGAGCGCGCGCGCGAATAAACCGTGTGAGAC 717
   |||||:|||||:|||||:|||||:|||||:|||||:
271 AlAthrAlaArgProThrSerLysProCysGlyCysLysProCysGlyPr 287
718 A 718
287 o 287
seq_name: p1r2:S59310

```

seq documentation block:
 Probable membrane protein YMR317w - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein YW924.09
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 29-Nov-1995 #sequence_revision 23-Feb-1996 #text_change 05-Nov-1999
 C:Accession: S59310
 R:Churcher, C.M.
 Submitted to the EMBL Data Library, September 1995
 A:Reference number: S59302
 A:Accession: S59310
 A:Molecule type: DNA
 A:Residues: 1-1104 <CHU>
 A:Cross-references: EMBL:Z54141; GSPDB:GN00013; MIPS:YMR317w
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: MIPS:YMR317w
 A:Map position: 13R

alignment_scores:
 Quality: 124.00 Length: 466
 Ratio: 0.588 Gaps: 21
 Percent Similarity: 45.279 Percent Identity: 23.176

alignment_block:
 US-09-303-518D-125/rev x S59310 ..

Align seg 1/1 to: S59310 from: 1 to: 1104

```

1331 TCATAGTTTCCAGACCTTTGCGCAACAGCGGCGGTATTCGATTTCGC 1282
   |||||:|||||:|||||:|||||:|||||:|||||:
350 SerSerThrSerSerValSerSerGluAlaProSerSer... 363
1281 CGGGCAGACGAGCTGCACAAACGAGCTCTTCGTCATTCGACG 1232
   |||||:|||||:|||||:|||||:|||||:|||||:
364 .....ThrSerSerValSerSerGlu 372
1231 AACCCATGCTGCGCGCTGCTGATGCGCGACGATTAAATCGGCAAA 1182
   |||||:|||||:|||||:|||||:|||||:|||||:
372 lAProSerThrSerSerValSerSerGluAlaProSerSerLys 388
1181 AGCAGGCTGGCAGCATATCCAGAGGCATACGCGCTCGTAATCAAT 1132
   |||||:|||||:|||||:|||||:|||||:|||||:
389 SerSerVal...MetSerSerGluValSerSerAla...Th 400
1131 CGGACCAATGCGCGCGCGCGCGCTGACGCGCTGTGAATGAGGA 1082
   |||||:|||||:|||||:|||||:|||||:|||||:
400 rSerSerLeuValSerSerGluAlaProSerAlaLysSerSerLeuAla 417
1081 GT...TTGTTTTCAGAAATGCGGAGGTTGTACGCGTGATG... 1041
   |||||:|||||:|||||:|||||:|||||:|||||:
417 rSerArgLeuPheSerSerLysAsnThrSerValThrSerThrLeuVal 433
1040 .....GAGTATTTTCGCGCTGCGCGCGCAACCCAGCGCAACGCTCTT 997
   |||||:|||||:|||||:|||||:|||||:|||||:
434 AlAthrGluAlaSerSerValThrSerSerLeuArgProSerSerGlu 450
996 GCTGCGGCTTCTGATACGGAATGATGATTGTAAGGTGCCAAT 947
   |||||:|||||:|||||:|||||:|||||:|||||:
450 rLeuAlaSerSerSerLeuIleGlu... 458
946 NATCGTGGCGCTTGTGTAATCGCGCGCTTCATACCGAACGGAAATC 897
   |||||:|||||:|||||:|||||:|||||:|||||:
458 ..... 458
896 AGCGGCTTGTCTGTACCAACATTCGCCGCGAGTAATTCGATCTTT 847
   |||||:|||||:|||||:|||||:|||||:|||||:
459 SerSerLeuSerThrGlyTyrAsnSerThrValSerThrThrThrSerAl 475
846 CGCACCAAAAGGTACGCAAGAGCGCGGTTTGTGACTGGAACACC 797
   |||||:|||||:|||||:|||||:|||||:|||||:
475 aAlaSerSerThrLeuGlySerLys... 483

```


[illegible]

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-772 <HA>
 A:Cross-references: GB:BA000007; PIDN:BA35761.1; PID:g13361805; GSPDB:GN00154
 A:Experimental source: strain 0157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: EGS2338
 C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] nc

Alignment_scores:
 Quality: 115.00 Length: 322
 Ratio: 0.728 Gaps: 12
 Percent Similarity: 49.068 Percent Identity: 22.050

Alignment_block:
 US-09-303-518D-125 x B90921 ..

Align seg 1/1 to: B90921 from: 1 to: 772

```

100 GAAGAAATATGCGCGTATGCGCCCTCGATGAAAGTCAAGAGAGCGATGC 149
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
48  LysGlnHisIleGlyAlaGlyGlyLeuGlyValSerValGlyAspGly 64
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
150 CGTCAAAAAGAGCGAGTCTGTTGAAGACAAAAGAAATCCGGCGCTGG 199
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
64  sValLeuArgGlyGlnProLeuThrArgGlyArgGlyLysMetLeuPro 81
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
200 TGTCTACTGCGCGCTTCAGGCAAAATCCCGCGATT..... 237
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
81  AlHis...AlaProThrSerGlyThrValThrAlaIleAlaProHisSer 96
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
238 .....CACGTGGCGGAAAGCGCGTACTTCAGTCTGATGTCGCGT 281
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
97  ThrAlaHisProSerAlaLeuAlaGlyLeu...SerValIleLeuAspAl 112
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
282 TGAAGGCAAGAC..... 294
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
112 aAspGlyGluAspGlyTrpIleProArgAspGlyTrpAlaAspTrpArgS 129
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
295 .....GAATCGAGTTTGAACGCTACCGACCTGACCGCTGACG 333
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129 eArgArgArgGlyGluGlyLeuIleGlyAlaGlyIleHisGlnPheGlyValAla 145
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
334 AACTTAAGCGCGGAGAGAGTGGCGCCCAACCTGATCCAAATCCGGTTTGG 383
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
146 GlyLeuGlyGly.....AlaGlyPhePr 153
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
384 GACTGGCGTGGCGACCGCTCCGTTTCAGCAAAATTCCTGCGCGATGCCG 433
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
153 oThrGlyVal.....LysLeuGlnGlyGlyAspL 164
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
434 AGCGGTTCCCATCTTCGTAATGCGAGTGCACCAATCCGCTGGCTGCC 483
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
164 yslIleGlyThrLeuIleIleAsnAlaIleGlyGlnProGlyTrpIleThr 180
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
484 GACCCATCGGTATTCATCAAGAGCGCGGAGTTTCAACGGCGGCT 533
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 AlaAspArgArgLeuMetClnAspGlyAlaIleGlyValGlyIle 197
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
534 GTTGGATTGAGCCGTTTGACCGAA..... 558
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
197 eArgGlyLeuAlaHisIleLeuGlnProArgGlyLeuLeuIleGlyIleG 214
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
559 .....CGCAAAATCCATGTTTGAAGCAGCTGGCGCGACAG 594
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
214 LuAspAsnLysProGlnAlaIleSerMetLeuArgAlaValLeuAlaAsp 230
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
595 GTGCGGTGTGAAGATGCTGCACATCGAAGACATCGAATTCGGCGGCGCC 644
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
231 .....SerHisAspIleSerMetLysValIleProThrLy 242
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
645 GCATCTTCGCGGT.....TTGAGTGGCAGCG 670
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  
```

```

242 sTrpProSerGlyGlyAlaLysGlnLeuThrTrpIleLeuThrGlyLysG 259
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
671 ACATTCATTTTCATCGAGCCGGTGGCGCGCAATTAACCGGTGGACCAAC 720
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
259 Inval.....ProHisGlyGlyArgSerSerAspIleGlyVal 271
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
721 AATTATTCAGATGTAAATTACC.....ATTGGCGGTTTGTTCGAAC 761
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
272 LeuMetClnAsnValGlyThrAlaTrpAlaValAlaValArgAlaValIleAs 288
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
762 AGCGCGTTCGACACCGAGCGGTGATTCGCCCTAGGTGTTCTCAAGTGA 811
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
288 pGlyGlnProIleThrGlyArgValAlaValThrLeuThrGlyGlyAlaIleA 305
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
812 ACAACCGCGCGCTTCGCTACCGCTTTCGCTGCGCAATGATTCGCAATT 861
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
305 laArgProGlyAsnValTrpAlaArgLeuGlyThrProValArgHisLeu 321
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
862 .....ACTGCGGCGCAATTGGTTGACACACACACCGCGTATTTCCGG 905
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
322 LeuAsnAspAlaGlyPheCysProSerAlaAspGlnMetValIleMetG 338
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
906 TTCGATATGAAACGC 921
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
338 yGlyProLeuMetGly 343
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  
```

seq_name: p1r2:T43481

seq_documentation_block:

Probable mucin DKFZp434C196.1 - human (fragment)
 N:Alternate names: protein DKFZp434B0635.1

C:Species: Homo sapiens (man)

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 20-Apr-2000

C:Accession: T43481; T34549; T17264

R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z22514

A:Accession: T43481

A:Molecule type: mRNA

A:Residues: 1-580 <AA>

A:Cross-references: EMBL:AL133561; NID:g6599133; PIDN:CAB63715.1; PID:g6599134

A:Experimental source: adult testis; clone DKFZp434C196

R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, October 1999

A:Reference number: Z21540

A:Accession: T34549

A:Molecule type: mRNA

A:Residues: 262-580 <POU1>

A:Cross-references: EMBL:AL122069; NID:g6102864; PIDN:CAB59245.2; PID:g7018420

A:Experimental source: adult testis; clone DKFZp434B0635

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18723

A:Accession: T17264

A:Molecule type: mRNA

A:Residues: 262-580 <POU2>

A:Cross-references: EMBL:AL117481; NID:g5911958; PIDN:CAB55954.1; PID:g5911959

A:Experimental source: adult testis; clone DKFZp434B061

C:Genetics:

A:Note: DKFZp434C196.1; DKFZp434B0635.1; DKFZp434B061.1

Alignment_scores:

Quality: 114.50 Length: 563
 Ratio: 0.540 Gaps: 25
 Percent Similarity: 37.635 Percent Identity: 21.492

Alignment_block:

US-09-303-518D-125/rev x T43481 ..

Align seg 1/1 to: T43481 from: 1 to: 580

```

1336 CCTTCTCAATGGTTTCAGACATTGGCAACGCGGCCGATTCGTAT 1287
|||||
29 ProSerArgAlaSerProThrArgMetProSerArgAlaSerLeuSme 45
1286 TTGGCCGGGCGAGAGAGCTGCACAAAG.....CGAGTCTCTTC 1246
: ||| ||| ||| |||||
45 tthProheArAlaSerLeuThrLysMetGluSerThrAlaLeu 62
1245 GTCCAAATTCGA..... 1235
|||
62 rgtHleuProArAlaSerLeuMetArghrProThrArgAlaSerLeu 78
1234 .....AG.1233
79 MetArghrProArAlaSerProThrArgLysProProArAlaSe 95
1233 CAACCCATGGCTGGCGGTATCGCCGAGTAAATGCGCA. 1184
|||
95 rProArghrProSerArghAlaSerProThrArgArghLeuProArghAla 112
1183 .....AAA 1181
112 ePromeGlySerProHisArghAlaSerProMetArghrProProArgh 128
1180 CGAGGCGGCGAGATTCGA.....GGCATCAGCGC 1147
|||||
129 AlaserProThrGlyThrProSerThrAlaSerProThrGlyThrProse 145
1146 CTCGTAAGTCAATCGCACCATTGCGCG.....TCGC 1112
|||||
145 rSerAlaSerProThrGlyThrProProArghAlaSerProThrGlyThr 162
1111 CGCGCTGACGCGTGTGAACCTGAAGAGTTGTTTTCAGCAATG 1062
|||||
162 rPro.....ArgAlaThr 166
1061 CGAGGCTGTACGCGTATGAGTATTGTCGCGCTGCGGCAACCCA 1012
|||||
167 .....AlaThrAr 169
1011 GCGCAACAGC.....TCTTTCGCGCGCTTCGTCAACGGAATCT 968
: ||| ||| ||| |||
169 gSerProSerThrAlaSerLeuThrArghrProSerArghAlaSerLeu 186
967 GATTGCGTACGCTGCCAATATGCTGCGCGCTGTGTATGCGCGCG 918
|||||
186 hrArghrPro.....ProProArghAlaSerPro 194
917 TTCATACGCAACGCG...GAAATCAGCGGTGTCT.....GTGTC 880
|||||
195 ThrArghrProProArghLysrProArghMetSerHisArghAlaSerPr 211
879 AACCAATGCGCGCAGTAAATTCGATCTTCGCACCAAAACGGTAC 830
|||||
211 ofhrArghrPro.....ProArghAlaSerProThrArgh 223
829 GCAAGAGCGCGCTTGTGACTTGAGAACCACTGACATTCACGCGC 780
|||
223 rgtProArghAlaSerProThrArghrProProArghLysrLeuArgh 239
779 TCGGTGTTCA.....CGGCTGTTCGAAA 754
|||
240 ThrSerHisArghAlaSerProThrArghMetProProArghAlaSerPro 256
753 CAAGCGCCAGTGAATTCATCTTGATTAATGATGTCACACGCGTT 704
|||||
256 rArghrPro.....ProA 261
703 TATTGCGCGCGCGCTCGATGAATGATGCTGCACCAACAGC 654
|||||
261 rghAlaSerProThrGlySerProPro...ArgAlaSerProMetThrPro 276
653 GCAGGATGCGCG.....CC 640

```

```

277 ProArghAlaSerProArghrProProArghAlaSerProThrThrThrPr 293
|||||
639 GCCGAATTCATGTTGTCGATGTTGGCAGCATTTTCAGCGCAGCTCG 590
|||||
293 oSerArghAlaSerLeuThrArghrProSerThrAlaSerProThrThrT 310
589 CGCCAGCTGCTTCAACATGATTTTGCTGCGGTCAGCAAGCGCTCAAT 540
|||||
310 hrProSerArghAlaSerLeuMetLysMetGluSerThrAlaSerLeu 326
539 ACCAAGCGCGCGCTTGAAA..... 519
327 ArghrProProArghAlaSerProThrGlyThrProSerArghAlaSerPr 343
519 ..... 519
343 oThrGlyThrProSerArghAlaSerLeuThrGlySerProSerArghAla 360
518 .....TCTCGCGCGCTTCTTTCGATTAATGACCGTAGGCTCG 483
|||||
360 eLeuThrGlyThrProSerArghAlaSerLeuThrGlyThrProSerArgh 376
483 GCAGCCAGCGGATTTGTCCTCATTCGATTCAGCAAGATGCGCAAGCTC 433
|||||
377 Alaser.....LeuThrGlyThrProSerArghAlaSerLeuThrGlyThr 391
432 GGCATTCAGCGCAGGATTTTCGTCGAGACGAGCGGTCGCGACGACATC 383
|||||
391 rProProArghAlaSerLeuThrGlyThrSerSerThrAlaSerLeuThra 408
382 ACAACCGGATTCGATTCGCGCGCACCTTCCTTCGCGCTTAAGTTT 333
|||||
408 rgtHrProSerArghAlaSerLeuThrArghrProse 420
332 GCCAGCGCTTACGTCGCTACGCTTCGAACCTCGATTCGTCGCTTC 283
421 .....SerSerLeuThrArghrProse 429
282 AACGGCATTCAGACTGACTGATGACGCGCTTTCGCGCAGGTGAATCG 233
|||||
429 rMetAlaSerLeu.....ThrArghrProProArgh.....A 440
232 CGCGCATTTTG.....CCTGAAGCGCGCGCATTAACACCAACGCGC 192
|||||
440 lArserLeuThrArghrProProArghAlaSerLeuThrArghrProPro 456
191 GGAATCTTTTGTCTCAACACGACTTGCGCTTTTTCGAGCGCATCGCC 142
|||||
457 ArgAlaSerLeuThrArghrProProArghAlaSerLeuThrArghrPr 473
141 TTCCTGACTTTCATGAGGCGCGCATACCGCATATTCCTCG..... 99
|||||
473 oSerMetValSerLeuLys...ArgSerProSerArghAlaSerLeuThra 489
98 .....CCAAAGCAACGCGACT.....TCGCTA 78
|||||
489 rgtHrProSerArghAlaSerLeuThrMetThrProSerArghAlaSerLeu 505
77 ATGGCGCGCGCGCTGTAACGCGCTTCGCGCTGCGC 39
506 ThrArghrProSerThrAlaSerLeuThrGlyThrPro 518
seq_name: p1r2:147141

```

seq documentation block:
 gastric mucin (clone PGM-2A) - pig (fragment)
 C/Species: Sus scrofa domestica (domestic pig)
 C/Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 03-Nov-2000
 C/Accession: 147141; 555315
 R/Turner, B.S.; Bhaskar, K.R.; Hadzopoulos-Ciadaras, M.; Specian, R.D.; Lamont, J.T.
 Gastroenterology 106, 200, 1994
 A/Title: Pig gastric mucin: isolation and characterization of a cDNA clone with a nov

A:Reference number: I47141; MUID:94102478
 A:Accession: I47141
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-528 <TR>
 A:Cross-references: EMBL:U10281; NID:9915205; PIDN:AAC48526.1; PID:9915208
 R:Turner, B.S.; Bhaskar, K.R.; Hadzopoulou-Cladaras, M.; Specian, R.D.; Lamont, J.T.
 Biochem. J. 308, 89-96, 1995
 A:Title: Isolation and characterization of cDNA clones encoding pig gastric mucin.
 A:Accession: S55315; MUID:95275264
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-528 <TR>
 A:Cross-references: GB:U10281; NID:9915205; PIDN:AAC48526.1; PID:9915208
 C:Superfamily: pig submaxillary mucin

Alignment_scores:
 Quality: 113.50 Length: 458
 Ratio: 0.493 Gaps: 19
 Percent Similarity: 50.218 Percent Identity: 20.961

alignment_block:

US-09-303-518D-125/rev x I47141 ..

Align seg 1/1 to: I47141 from: 1 to: 528

```

1343 CAGCCTTCCTTCGCAATG.....GTTTCAGACATTTCGCGCAACAG 1303
149 GlnProSerSerSerSerSerProProIleSerSerThrVal..... 162
1302 CGGGCCGATTCGATTTTGGCCGGGCGAGAGACAGCTGCACAAAGCGAGT 1253
163 .....SerValGlnProSerSerSerSer.....SerAlaProI 174
1252 CTTCTGCTGCATTCGCAAGCAACCAATGCTTCGCGCGCTG...TCGGTA 1206
174 hrThrSerAlaThrSerValGlnProSerSerSerSerSerProIle 190
1205 TCGCCGACGATTAATCGCCCAAGAGAGGCGGCGAGATATCCAAAGG 1156
191 SerSerThrValSerValGlnThrSerSerSerSerSerSerValProI 207
1155 CAGCAGCGCTGTAAGTACCAATCGGACCAATGCGCGCGCTGCGCCCG 1106
207 rSerThrThrSerValGlnPro.....SerSerSerSerSerValProI 222
1105 TGACGCTGTGTGAACCTGAAGAGTTTGTTCAGGAATGCGCGAGG 1056
222 hrThrSerAlaThrSerValArgSerSerSerSerSerSerThrPro... 237
1055 GTTGTACGCGCTGATGAGATTTGTCCGCTGCGGCGCAACCGCGGAA 1006
238 .....IleProSerThrThrSerValGlnProSe 247
1005 CAGCTCTTTCGTCGCGCT...TCTTCGATTAACGAATCTGATGTGTG 959
247 rSerSerSerSerAlaProIhrThrSerAlaThrSerVal..... 260
958 AGGTCGCCAATATATGTCGCGCCTTGTGTAATCGCGCGCTCAATACC 909
261 ..GlnProSerSerSerSerSerThrProIleProSerThrThrSerVal 276
908 GAACCGGAATACAGCGGTTGTGTGTAACCAATTCGCCCGCGAGTAA 859
277 GlnPro.....SerSerSerSerSerAlaProIhrThrSe 288
858 TTGCGATCTTTCGACCCCAACGATACGACAGAGCGCGGTTGTGA 809
288 rAlaThrSerValGlnProSerSer.....S 297
808 CTTGAGAACCACTAGGCAATACAGCGCTCGGTTCAGACGCGCTGTT 759

```

```

297 ..... 297
erSerSerProIleSerSerThrIleSerVal.....GlnProSer 311
758 GCAACCAAGCGCCCAATGTAATTCATCTGTAATGATGATGTCACAC 709
312 SerSerSerSerSerProIhrThrSerThrThrSerValGlnProSer 328
708 GGTTCCTTTCGCGCCGACCGCGCTCGATGAATGAATGCGCGCACTCA 659
328 rSerGlySerAlaProIhrThrSer.....AlaThrSerValG 341
658 AACCGCAGATGCGGCGCGCAATTCATGTTTCGATGTTGCGACGA 609
341 InProSerSerSerSerSerProIleSerSerThrIleSerValGln 357
608 TTTTCAGACGCGACGTCGTGCGCAGCTTACAAACATGATGATTTGCG 559
358 ProSerSerSerSerSerSerProIhrThrSerThrThrSerVal.... 372
558 TTGCGTCAAAGCGCTCAATACCAAGCGCGCTTGAATCCTCGCGC. 510
373 .....GlnProSerSerSerGlySerAlaP 381
509 .....GTTCTTTGATTAATACCGTAGGCTGCGGACCGCGATGTTG 465
381 rThrThrSerAlaThrSerValGlnProSerSerSerSerSerValPro 397
464 TCCATGCGATTGACGACATGCGCAACGCGTCGCGATCGACGCGAGAT 415
398 ThrThrSerAlaThrSerValArgSerSerSerSerSerSerThrProI 414
414 TTTGTCGACGAGCGGTCGCGACGCGAGTCCCAAAACCGGATTCGATCA 365
414 eProIhrThrThrSerValGlnProSer..... 423
364 GGTTCGCGCGCACCTTCCTTCGCGCTTAAGTTGCCAGCGCTTCAGTTCG 315
424 .....SerSerSerSerValProIhrThrSerAlaThrSerVal 436
314 TAGCGTCAAACTGATTTGCTGCG...TTGCTTCACAGCGCAATCAGCAC 268
437 GlnThrSerSerSerSerSerThrProIleProSerThrThrThrSerValG 453
267 TGACTGAAGTACGCGCTTTTCGCCACGCGTGAATGCGCGATTTTTCGCTG 218
453 nProSerSerSerSerSerAlaPro..... 461
217 AAGCGGCGCAGTAACACACGCGCGGATTCCTTTGTCTTCAACAGC 168
462 .....ThrThr.....SerAlaThrSer 467
167 ACTTCGCTTTTTCAGCGCAATGCGCTTCCTTACCTTCATTCAGCGGCGG 118
468 ValGlnProSerSerSerSerSerSerProIleSerSerThrIleSerVal 484
117 CATACCGCATATTCCTGCGCAACGACGACTTCGATATG..... 75
484 GlnProSerSerSerSerSerProIhrThrSerThrThrSerValG 501
74 .....GCCGCGCGCTGCTGAACGCTTCTGCTCGGCTG 42
501 InProSerSerSerGlySerAlaProIhrThrSerAlaThrSerValGln 517
41 CCGCGCATGCGCAGGTTTACAGCT 18
518 ProSerSerSerSerSerProPro 525

```

seq_name: p1r2:F75196

seq_documentation_block:
 pyruvate, water dikinase (EC 2.7.9.2) PAB0057 - Pyrococcus abyssi (strain Orsay)
 N:Alternate names: phosphoenolpyruvate synthase
 C:Species: Pyrococcus abyssi

C>Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 18-Nov-1994
 C/Accession: C39135
 R:Holmes, M.L.; Dyall-Smith, M.L.
 J. Bacteriol. 173, 642-648, 1991
 A:Title: Mutations in DNA gyrase result in novobiocin resistance in halophilic archaea
 A:Reference numbers: A39135; MUID:91100352
 A/Accession: C39135
 A>Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-437 <HOL>
 A/Cross-references: GB:M38373

alignment_scores:
 Quality: 112.00 Length: 450
 Ratio: 0.626 Gaps: 23
 Percent Similarity: 39.778 Percent Identity: 24.222

alignment_block:
 US-09-303-518D-125 x C39135 ..

Align seg 1/1 to: C39135 from: 1 to: 437

```

36 CGCGGGACAGCGGACGCGT.....TTACGACGGCC 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
94 ArgAspGlnAspGlnAlaProArgArgHisLeuArgArgArgVa 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
71 CGCGCATTCGGAAGCGCGTTCGCGAAGATATGCGGATGCGC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
110 GlnGlnHisArgGlnArg..... 116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 CCCGCGAAGATCAAGAGAGGATGCGGTCAAAAAGCCAGTGT 170
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
117 .....AsnArgHisAlaGlnAspArgArgAlaProGlnValAspSerArg 131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
171 GTTGAAGACAAAAGATCGCGCGTGTGTTTACTGCGCGCC 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 LeuArgGlnGlnHisGlnHisProArgGlnArgHisAlaSerAspArgVa 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 .....TTCAAGCAAAATCGCGCGATTCACCGTGGC 246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 GlnAspGlnAlaHisProArgArgGlnArgGlnGlnProArgH 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 GAAAGCGCGTACTTCAGTCAGTCGATGTCGCGTGAAGGCA 290
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
165 AlaArgGlnArgPro.....ArgArgArgGlnProPro 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 CGAGCAATGAGATTGAACGCTACGACCTGAGCGCTGCAAACTAA 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
176 ArgArgGlnArg.....SerArgGlnHisArg 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
341 GCGGCGA...AGAAGTGGCGCGACCTGATCCATCCGCTTTGGAGCT 387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
185 GlnGlnHisLeuArgGlnAlaProArgPro...AlaValArgGlnProAsp 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
388 GCGCTCGGACCGCGTTCAGCAAAATTCGCGCGATGCGGAGCC 437
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 .....GlnAspGlnAlaArgGlnLeuArgGln 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
438 GTTCGGCAT.....CTTCG 451
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 ProArgHisArgArgGlnArgHisProProArgAlaArgAspValLeuArg 226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
452 TCAATGCGATGACACCAATCGCGTGGCGCGACCTGCTATTC 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
226 GlnGlnArgProGlnHisGlnAsp.....GlnHisHisL 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
502 AAAAGCGCGCGAGATTCAAACGCGCGCTGTATGAGCGGCTT 551
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
237 euGlnGlnArgArgGln...ArgProArgProGlnGlnArgGlnAlaGly 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
552 GACCGAGCGCAAAATTCATGTTGTAAGGCGAGCTGGCGACGTCGCCGT 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

253 ArgGlnAlaHisProProGlnVal.....ArgAlaArgIleu 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
602 CTGAAATGCTGCCACATCATGAAACATGATTCGGCGCGCGATCC 651
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265 rLeuAlaAlaGlnGlnAlaArgGlnLeuProGlnProArgGlnGly 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
652 GCGGCTTGGATGCGACGCGCATTCATTTCATGACCGCGTGGCGGAA 701
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
282 AlaArg.....ThrValHisArgGlnGlnArgGln 292
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
702 TAAACCGTGGACCATCATGATGATGATTAATTACATGGCGCT 751
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
293 ...GlnArgValGlnGlnAlaGlnProArgProGlnValProGlnAsp 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
752 TGTTCGACACGCGCGCTGTAACACGCGCGGTGATTCCTAGGTGT 801
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
308 eAlaProGlnGlnGlnAspSerGlnArgGlnArgGlnProProArgProH 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
802 TCTCAAGTCACAAACCGCGCTTCGCTACCGTTTGGTGCGCAAGT 851
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
325 AsSer..... 326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
852 ATCGCAAAATTCGCGCGCAATGTTGACACAGAACCGCTGATTT 901
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
327 .....ArgGlnArgArgAsp.. 331
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
902 CGGCTTCGATTCGAACGCGCGATTCACAAAGCGCGACGATTTTGG 951
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
332 ...ThrGlnAlaHisArgHisArgHisArgArg..... 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
952 GAGCGCTACCAATCAGATTCGTTATGCAAGAGCGCGAGCAAGA 1001
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
343 .....ArgArgArgValArgHisArg 349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1002 GCTGTTCGCGCTGGTTCGCGCGACCGCAAAATTCATCAGCGCTA 1051
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
350 GlnGlnAlaLeu..... 353
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1052 CAACCTCGCGCATTTCTGAAAAAACAACCTTCAGTTCAACACAGCC 1101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
354 ...ProAlaAlaHisProAspAspArg.....A 362
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1102 GTCAAGCGCGCGACCGCGCGATGTCGATGTTACTTACGAGCGGT 1151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 rArgArgArgArgArgAlaHis.....Pro 370
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1152 GATGCCCTTGATATCCCTGCCACCGCTCTTTCGCGGATTAATCGTCG 1201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
371 AspAlaAlaAlaValAlaSerValProAlaHisAlaProAlaHisArgG 387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1202 GCGATACCGACAGCGCGAGCATTTGGTTCCTGGAATGGAGAGAGA 1251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
387 rArgGlnArgValArgGlnSerThrAla.....AlaValProArgProL 402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1252 GACCTCGCTTTCGACGCTTCGTCGCCGCGCAATGCAAAATGCGGCC 1301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
402 euProArg.....GlnHisLeuArgArgArgGlnArgGlnAlaGlyPro 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

seq_name: p1r2:152257

seq_documentation_block:
 epistatin - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
 C/Accession: I52257; 165210
 R:Yos, H.L.; De Vries, Y.; Hilkens, J.
 Biochem. Biophys. Res. Commun. 181, 121-130, 1991
 A:Title: The mouse epistatin (Muc1) gene and its promoter. Rapid evolution of the rpr
 A:Reference number: I52257; MUID:92068178
 A/Accession: I52257
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-631 <RES>

A:Cross-references: GB:M77226; NID:g199835; PIDN:AAA39754.1; PID:g199837
 A:Accession: I65210
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-631 <RE2>
 A:Cross-references: GB:M84683; NID:g199842; PIDN:AAA39756.1; PID:g199843
 A:Gene: Muc1
 A:Introns: 20/1; 454/3; 472/2; 517/1; 557/3; 607/3

alignment_scores:
 Quality: 112.00 Length: 445
 Ratio: 0.599 Gaps: 19
 Percent Similarity: 42.022 Percent Identity: 23.371

alignment_block:

US-09-303-518d-125/rev x I52257 ..

Align seg 1/1 to: I52257 from: 1 to: 631

```

1337 TCCCTTCATAGTTTCCAGCATTTG.....CGCAACAGCGG 1300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
38 SerSerSerLeuAlaSerThrThrProValHisSerSerAsnSerAs 54
1299 GCCGATTCGTATTGTCGCCGGGAGAGAGCTGCACAAACGAGTCTT 1250
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
54 ProAlaThrArgProGlyAspSer..... 63
1249 CTTCGTCATTCACAGCAACCAATGCCTGCGCTGCGTATCGCGG 1200
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 .....ThrSerSerPro 67
1199 ACGATTAAATCGCGCAAAAGCGGTGGAGATATCCAAAGGCGATCA 1150
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 ValGlnSerSerThrSerProAlaThrArgAlaProGlyAspSerTh 84
1149 GCGCTCGTAGTACCAATCGGACCATGGCGGCGCGCGCGCTGACGG 1100
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 rSerThrAlaValLeuSerGlyThr.....SerSerProAlaThrThi 99
1099 CTGTGTGAACCTGAAGAGTTGTTTTCAGGAATGGCGGAGGTTGTA 1050
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 LeProValAsnSerAlaSer..... 105
1049 CCGGTGATGAGATATTGTCCGGCTGGCGGCGCAACCGGACAGCTC 1000
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
106 .....SerProValAlaHisGlyAspThrSerSerProAlaThrSe 119
999 TTGTCGCGGCGCTTTCGATTAAGGAATCTGATTGTGTGACGTCGA 950
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
119 rLeuSerLysAspSer..... 124
949 AATAATCGTGGCGGCTGTGTATCGCGCGCTTCAATACCGAACCGGAA 900
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
125 .....AsnSerSerProValAlaHisSerGlyThrSerSerAlaProAla 139
899 ATACGCGGCTGTCTGTCAACCAATTCGCCGCGAGTAATTGGCATAC 850
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
140 ThrThrAlaProValAspSerThrSerSerProValAlaHisGlyThr 156
849 TTTCGCAACCAAAAGGTACGAGGAGGCGGTTGTTGTAATTCGATCA 800
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 rSerSerProAlaThr.....SerProProGlyAspSerThrSerSerP 171
799 CACCTAGGCAATCACGCGCTCGGTGTACAGACGCGCTGTGCAACAAA 750
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
171 rAspHisSerSerThrSerSer.....ProAlaThrArg 182
749 CGGCAATGTAATTCATCTTGAATATGATGTCACACGGGTTTATT 700
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
183 AlaProGlyAspSerThrSer.....ThrAlaValLeu 193

```

```

699 CGGCGGACCGGCTCGATGAATGATGCTGCGCACTCAACCGGACG 650
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
193 rSerGlyThrSerSer.....ProAlaThr 201
649 GATGCGGCGCGCGAATTCATGTGTTGATGTGGCAGCATTTTCAAC 600
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 hThrAlaProValAspSerThrSerSerProValAla.....HisAsp 215
599 GGCAGCTCTGCGGCGACCTCCCTTACAAACATGATTTTGGTGGTCA 550
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 AspThrSerSerProAlaThr..... 222
549 ACGCTCAATACCAACAGCGCGCTTGAATCCTCGCGGCTCTTTGA 500
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
223 .....SerLeuSerGlyAspSerAlaSerSerP 232
499 TAATGACCGTAGGGTCGCGACCGCAGCGGATGGTGTCCATCCGATTGACG 450
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 roValAlaHisGly..... 236
449 AAGATGCGCAACGCGCTCGCATGACGCGAGAAATTTGCTGACGAGCG 400
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
237 .....GlyThrSerSerProAlaThrSerProLeuArgAspSe 249
399 GTGCGGACGCGCAGTCCACAAACCGGATTCGATCAGTTGCGCGGACCT 350
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
249 rThrSerSerProValHisSerSerAlaSerIleGlnAsnIleGlyThr 266
349 CTTCGCGCGCTTAAAGTTTGCACGCGCTTCAGGTGCGTAAGCGTTAACTCG 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 hSerAspLeu..... 269
299 ATTTCGTCGTTGCT.....TCACGCAATCAGCAGTGTGAAG 259
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
270 .....AlaSerThrProAspHisAsnGlyThrSerValThrThrThrSerSe 285
258 TACGCGCTTTTCGCGCACGCGTGAATCGCGGATTTTGCCTGAACCGCGG 209
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
285 rAlaLeuGlySer.....AlaThrSerProAspHisSer 297
208 CAGTAACACCGACCGCGGATCTTTTGTGCTTCAACACAGCTGTGGCT 159
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 LyrThrSerThrThr.....ThrAsnSerSerGlySer 307
158 TTTTGAAGCGCATGCGCTTCCTGACTTTCATCGAGGCGGCGATACCGGC 109
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
308 ValLeuAlaThrThr.....ProVal 314
108 ATATTCCTTTCG...CCAAAGCAACGCGACTTCGCTAATGCGCGGG..... 69
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 LyrSerSerMetProPheSerThrThrLysValThrSerLysSerAla 331
68 .....CGGTGCTAAACGCGCTCGCTCGGCTGCGCC 39
331 leIleProAspHisAsnLysSerSerValLeuPro 342
seq_name: p1r2:C75580

```

seq_documentation_block:
 adenine deaminase-related protein - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: C75580
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.U.
 M.; Shen, M.; Yamathavan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioreistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: C75580
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-376 <WHI>


```

282 .....TGAGGCAAGCAAGATTCGATTTGAACSTACGACCTG 322
      |||||.....|||
132 G1yAlLeuArGlnIsArGlnIsArG..... 141
      |||||.....|||
323 AAGGCGTGGCAAACTTAAGCG.....CGAAGAATGGCGCGGCAACCTG 366
      |||||.....|||
142 .AlArGGLyGlnArGlnArGlnArGlnArGlnArGlnArGln 158
      |||||.....|||
367 ATCCAAATCCGTTTGTGGACTGCGCTGGACCCGCTGCTGCAAAAT 416
      |||||.....|||
158 IsPrArGGLyArGlnArGlnArGlnArGlnArGlnArGlnArGln 173
      |||||.....|||
417 TCTGCGCTGATGCGGACG.....GTTCG 442
      |||||.....|||
174 .....ArGArGArGArGArGArGlnArGlnArGlnArGlnAr 188
      |||||.....|||
443 CCATCTGTGTCATGCGGATGGACACCAATCCGCTGCTGCGGACCC 488
      |||||.....|||
188 GnlS.....G1yArArGlnIsArGlnAlAsPrArGArGPrArGAr 203
      |||||.....|||
489 .....TACGTCATTA.....CAAGA 506
      |||||.....|||
203 sPrArGArGAlAg1ylnIsArGlnArGlnArGlnArGlnArGln 219
      |||||.....|||
507 AGCCGC.....CGAGATTTCAAACGCGCGCTGTGGATTGA 544
      |||||.....|||
220 AlArArGPrArGlnArGlnArGlnArGlnArGlnArGlnArGln 234
      |||||.....|||
545 GCCGTTTGGACGCAAGCAAAATCCATGTTTGAAGCAGCTGGCCG 590
      |||||.....|||
235 .PrAlArGlnArG.....HisPrAlArGlnArGlnArGlnArGln 249
      |||||.....|||
591 .....AGACGTGCGCTGAAAAATGC 611
      |||||.....|||
249 rGAlArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 265
      |||||.....|||
612 TGCCAAATCCGCAAAATTCGATGGCGCGCGCTGCTGCGGTTGA 661
      |||||.....|||
266 LeuGlnArGArG.....G1yArGlnArGlnArGlnArGlnArGln 279
      |||||.....|||
662 GTGGGACGCAATTCATTCGACGCGT.....CGCG 696
      |||||.....|||
279 nG1yArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 296
      |||||.....|||
697 GCGAATTAACCGCTGGACATTCATTCAGATGATTAACATTCG 746
      |||||.....|||
296 rGln.....AsPrArGlnArGlnArGlnArGlnArGln 304
      |||||.....|||
747 CCG.....TT 751
      |||||.....|||
305 PrArGPrArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 321
      |||||.....|||
752 TGTTCGAAACGCGCTGCAACCGCGCGCTGATTCGCGTACG.... 797
      |||||.....|||
321 yAlArGlnArGPrArGlnArGlnArGlnArGlnArGlnArGlnArGln 338
      |||||.....|||
798 .....TGTTCTCAAGTCAGCAAAACGCGCTGCTGCTGCGTACG 836
      |||||.....|||
338 lAsPrArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 354
      |||||.....|||
837 TTTGGTGGCAAGATTCGCAAAATTCGCGCGGCAATTCGTTGACAG 886
      |||||.....|||
355 GlnG1yArGln.....G1yAlArGlnArGlnArGlnArGlnArGln 365
      |||||.....|||
887 ACAACCGCGTATTCGCGCTGCGTATTCG.....CGCGCGGATTACA 930
      |||||.....|||
365 gnlSPrArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 382
      |||||.....|||
931 CAAGCGCGCAAGATTCGCGCGGCGCTGACCAATCAGATTCGCGTAT 980
      |||||.....|||
382 euArGlnArGln.....PrArGlnArGlnArGlnArGlnArGln 391
      |||||.....|||
981 CGAAGAGCGCGCAAGAGCTGTTCGCGTGGCTGCGCGCGCGCGG 1030
      |||||.....|||

```

```

      |||||.....|||
392 ArGArGGLyPrArGlnArGln.....G1yAlArGlnArGln 404
      |||||.....|||
1031 ACAAAATCTTCATCAGCGCTGACACCGCTGCGGCAATTCCTGAAACAA 1080
      |||||.....|||
404 yGlnG1yArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 420
      |||||.....|||
1081 CTCTTCAAGTTCACACCGCTGACACCGCGCGCTGCGGCGCTGCGTGC 1130
      |||||.....|||
421 LeuAlArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 437
      |||||.....|||
1131 GATTGGTACTTACGACGCGCTGATGCGCTGATATCTGCGGCGCGCTG 1180
      |||||.....|||
437 uTrMetThrSerG1yArGln.....LeuArGlnArGlnArGlnArGln 452
      |||||.....|||
1181 TTTTGGCGCGATTATTCGTCGCGGATTCGAC..... 1212
      |||||.....|||
452 euCyArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 1250
      |||||.....|||
1213 .....AGCGCGACGCGATTCGCTGCTGCAATTCGACGAGCA 1250
      |||||.....|||
469 AsPrAlArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 484
      |||||.....|||
1251 AGACCTGCTTTCGACGCTGCTGCGCGCGCGCAATTCGATTCGCGC 1300
      |||||.....|||
484 euArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 500
      |||||.....|||
1301 CGCGTTCGCGCAAGCTGCGCA 1324
      |||||.....|||
501 SerThrSerArGlnArGlnArGlnArGlnArGlnArGlnArGln 508
      |||||.....|||
seq_name: p1r2:T43311
seq_documentation_block:
fatty-acyl-CoA synthase (FC 2.3.1.86) beta chain - fission yeast (Schizosaccharomyces
N/Alternate names: fatty acid synthase beta chain
C/Species: Schizosaccharomyces pombe
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C/Accession: T43311
R/Niwa, H.; Katayama, E.; Yanagida, M.; Morikawa, K.
Protein Expr. Purif. 13, 403-413, 1998
A/Title: Cloning of the fatty acid synthetase beta subunit from fission yeast, coexpr
A/Reference number: Z22415; M01D:98359895
A/Accession: T43311
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2073 <NTM>
A/Cross-references: EMBL:AB010274; NID:94115479; PID:BAA36384.1; PID:94115480
C/Genetics:
A:Gene: fast1
C:Superfamily: yeast fatty-acyl-CoA synthase beta chain
C/Keywords: acetyl-CoA; acyltransferase; carbon-oxygen lyase; coenzyme A; fatty acid
alignment_scores:
Quality: 110.50 Length: 465
Ratio: 0.521 Gaps: 25
Percent Similarity: 45.591 Percent Identity: 20.215
alignment_block:
US-09-303-518D-125 x T43311 ..
Align seg 1/1 to: T43311 from: 1 to: 2073
217 TCAGCAAAATTCGCGCGATTCACCGTGGCAAAAGCGCGTATTCAGTC 266
      |||||.....|||
948 ThrG1yArGln.....MetArGlnArGlnArGlnArGlnArGlnArGln 962
      |||||.....|||
267 AGTC.....GTGATTGCGCTTGAAGCAAGCAAGCAAAATTCAG..... 303
      |||||.....|||
962 pValG1yArGlnArGlnArGlnArGlnArGlnArGlnArGlnArGln 979
      |||||.....|||
304 .....TTGAACGCTACGACCTGAAAGCGCGTGC 333
      |||||.....|||

```

```

979 yrrProValAlaAlaArgPheLeuAspAlaTyrProGluAlaSerThrGln 995
334 AACCTAAGCGGCGAAGAGAGTGGCGCAACGATCAATCCGCTTGTG 383
996 AspLeuAsnThrGlnAspAlaGln.....PhePh 1005
384 GACCTCGCTGGCGCACCGCTCGCTGACGAA.....ATTCCTG 421
1005 eYrrSerLeuYssSerAsnProPheGlnLysProValProPheThrPro 1022
422 CCGTGCAT..... 429
1022 LaileAspAspThrPheGluPheTyrPheLysAspSerLeuTrpGln 1038
430 GCGGAGCGGTTGGCCATCTTCGTCAT.....GGAT 461
1039 SerGlnAspLeuAlaAlaValAlaGlyGlnAspValGlyAlaValAlaI 1055
462 GGACACCAATCCGCTGGCTGGCGACCGTACGGTCATTATCAAGACCG 511
1055 eLeuGlnLysProMetAlaAlaLysHisSerThrLysValAsnGluPro 1072
512 CCGAGATTTTCAAGCGCGCTGTGGATTGAGCGCTTGACCGACGC 561
1072 LatsyGlu.....LeuLeuAspGlyIleAsnGluThr 1082
562 AAAATC...CATGTTTGTAAAGCA.....CTGGCGCAGAC...GT 596
1083 HisIleGlnHisPheIleLysLysPheTyrAlaGlyAspGluLysIle 1099
597 GCGCTGTGAATAATGCTGCCAACAATCGAACAACATGAATTCGGCGCCGC 646
1099 eProIle.....ValGluTyrPheGlyGlyValP 1109
647 ATCTGCGCGTTTCAT..... 663
1109 ropProValAsnValSerHisLysSerLeuGlnSerValSerValThrGlu 1125
664 .....GGCAGCGCAATTCATTTCATCGACCGCGCGCGCAATTAAC 707
1126 GluAlaGlySerLysValTyrLysLeuProGluIleGlySerAsnSerAl 1142
708 CGTG..... 711
1142 AlaupProSerLysLysLeuThrPheGluLeuAlaGlyProGluTyrT 1159
712 .....TGGACC 717
1159 hrrTrpPheArgAlaIlePheThrThrGlnArgValAlaLysGlyTrpLys 1175
718 ATCAATTATCAAGATGATATACATGCGCTTGTTCACAACAGCGCG 767
1176 LeuGlnHisAsnProVal.....ArgArgIlePheAlaPro.....1187
768 TCTGAACACCGAGCGCGTATGTCCTAGGTGTTCTCAACGACAACAAC 817
1188 ArgTyrGlyGlnArgAlaValValLysGlyLysAspAsnAspTrpVal 1204
818 CCGCGCTCTGGGTAC.....GTTTGGGTGCGAAGTA 852
1204 aGluLeuLysGluThrGlnSerGlyAsnTyrValLeuAlaAlaArgLeu 1220
853 TCGCAAAATTAATGCGCGAATGTGTGACACAGACACCGCGATTTTC 902
1221 Ser.....TyrAspGlyGluThrIleValValSe 1230
903 CGGTTCGATTGAAGCGCGGATTAACAAGCGCGCAC...GATTATT 949
1230 lmetPheGlnAsnArgAsnAlaLeuLysGlyValHisLeuAsnPheI 1247
950 TGGGACGCTAC.....CACATTCGATTTCCGTTATCGAA 984

```

```

1247 eurrPheLysTyrGluProSerAlaGlyTyrSerProValSerGluIleLeu 1263
985 GAAGGCCGACGAAGAAGCTG.....TTCGGCTGGGT..... 1017
1264 AspGluArgAsnAspArgIleLysHisPheTyrTrpAlaLeuThrPheG 1280
1018 .....GGCGCGCGACGACAAATACATCCATACCGCTAACACCGCGCC 1063
1280 yGluGluProTyrProGluAsnAlaSerIleThrAspThrPheThrGlyP 1297
1064 ATTTCCTGAAA.....AACAACTCTCAAGTTCAACACAGCGCTC 1104
1297 roGluValThrValThrGlyAsnMetIleGluAspPheLysArgThrVal 1313
1105 AACGGCGCGACCGCGCATGGTGGATGGTGTACTTACAGCGCGGTAT 1154
1314 GlyAsnHisAsnGluAlaTyrThrLysArgAlaIleArgLysArgMetAl 1330
1155 GCGCTTGAT..... 1165
1330 aProMetAspPheAlaIleValAlaGlyTrpGlnAlaIleThrLysAlaI 1347
1166 TCTTGGCCACCGCTGTTTGGCGATTTA.....ATGTCGCGCATAC 1209
1347 lePheProLysAlaIleAspGlyAspLeuLeuArgLeuValHisLeuSer 1363
1210 GACAGCGCGCAGCATTTGGGTTCCTTGGAAATTCAGACGAAGAAGAC 1254
1364 AsnSerPheArgMetValGlySerHisSerLeuMetGluGlyAsp 1378

```

seq_name: p1r2:T39207

```

seq_documentation_block:
fatty acid synthase, subunit beta - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C/Accession: T39207
R/Wood, V.; Rajandream, M.A.; Barrell, B.G.; Ransberger, U.; Pohl, T.
submitted to the EMBL Data Library, September 1999
A/Reference number: Z21835
A/Accession: T39207
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2073 <WOO>
A/Cross-references: EMBL:AL110469; P1DN:CAP5157.1; GSPDB:GN00066; SPDB:SPAC926.09c
A/Experimental source: strain 97zh-, cosmid c926
C/Genetics:
A/Gene: SPDB:SPAC926.09c
A/Map position: 1
C/Superfamily: yeast fatty-acyl-CoA synthase beta chain

```

```

alignment_scores:
Quality: 110.50 Length: 465
Ratio: 0.521 Gaps: 25
Percent Similarity: 45.591 Percent Identity: 20.215

```

alignment_block:

US-09-303-518D-125 x T39207 ..

```

Align seg 1/1 to: T39207 from: 1 to: 2073
217 TCGAGCAAAATTCGCGGATTCACCGTGGCAAAACGCGTACTTCAGTC 266
948 ThrGlyLysLeu.....MetArgGlyIleGluArgPheThrGlnAs 962
267 AGTC.....GTGATTCGCTTGAAGCAACGACGAATTCGAG..... 303
962 pValGlyLysThrThrLeuIleGlnAsnPheGlnAspLeuAsnAspProt 979
304 .....TTGACGCTACGACCACTGAACGCGTGGCA 333
979 yrrProValAlaAlaArgPheLeuAspAlaTyrProGluAlaSerThrGln 995

```



```

139 GAAGGAGTCCGTCACCAAGCCAGTCTTTGAAGACAAAAGA 188
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
864 ValAlaIleAlaLeuArgLarGSerLeuThrValThrProLysAsnAs 880
189 TCCGGCGCGTGTACT.....G 208
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
880 mdLynIleAlaLeuArgLarGSerAspAlaCysLeuValGlyIleHis 897
209 GCGCGCTTCAGGCAAAATCGCGCGATTCACCGTGGGCAAAAGCGCTA 258
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
897 LaAspAlaSerValArgLarGLeuAlaSerIleHis..... 907
259 CTTCAGTACGTGATT.....GCCGTGAAGCAAGCAGA 296
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
908 .....SerLeuAlaIleAspAlaProGlnPheValGlnIleGlyGlnI 922
297 AATCAGATTGACGCTACGACCTGAAGCGCTGGCAAAATTAAAGCGCG 346
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
922 ValAlaGlnValIleLeuAlaGlnAspGlu...ThrClyAlaSerPheG 938
347 AAGAAGTCCGCGCAACCTGATCCATCCGTTGTGACGCGCTGCGC 396
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
938 IuLysGlnHisArgProLeuAlaAspAlaGlnLeuAspAlaSerAsn 954
397 ACCGCTCCGTCAGCAAAATTCCTGCGCTGAGTCCGAGCGCTGGCCAT 446
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
955 HisValIleLeuThrLys.....ValAspIleuArgLarGlyThrLe 968
447 CTTCGATATGCGATGACACCAATCCGCGCGCGCGACCCACAGGTCA 496
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
968 uArgAlaAsnSerIleGlyThrValSerLeuSerAla.SerSerLysSer 984
497 TTATCAAGAAGACCGCGAGCATTTCAACGCGCGCTGTGTATTGAGC 546
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
985 SerSer.....GlyArgValLeuSerSerArg..... 993
547 CGTTTGACGACGCAAAATTCATGTTTGAAGCAGCTGGCGACACGT 596
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
994 .....ProHisThrValGlnIlePheSerProIlePheLeuGlnProL 1008
597 GCCGTCTGAATGCTGCCAATCGAATCGAATCGAATCGAATCGAAT 639
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1008 yArGLeuThrLeuIleProAspSerLysPheGlnLeuGlnValAlaGly 1024
640 GCGCGCGATCTGCGC.....GCTTTCAGTGGCAGCAGCAGCAT 674
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1025 GlyProGlnProThrProLeuAspPheSerLeuAsnAsnSerMetIle 1041
675 TCATTCATCGAGCGC.....GTGCGGCGCGA 700
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1041 eAlaSerIleGlnProAsnAlaLeuIleThrSerSerGlnLeuGlyTyr 1058
701 ATTAACCGCTGGACCATTAATCAAGAT..... 732
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1058 hValAlaIleThrGlyThrValArgValGlyAspGlyHisValThrLeuAsp 1074
733 .....GTAATACCATGCGCTGTTGTTGGACACA..... 762
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1075 ThrValValLeuArgValAlaSerLeuGlyGlyIleIleLeuSerAlaSe 1091
763 .....GCCGCTCTGAC..... 774
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1091 rSerArgLysValGlnThrGlyArgValAsnLeuArgLeuArgGly 1108
775 .....ACGAGCGCGATGTCCTTAGTGCTCT... 804
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1108 alIleAlaGlyAlaGlnAspGlnProPheAlaPheGlyAlaIle 1124
805 .....CAAGTCAACAACCGCGCTCTGCGGAC 833
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1125 TyrProPheLysValThrTrpSerValSerProSerValLeuPheTh 1141

```

```

834 CGTT.....TGGGTCGAAATATCGCAAAATTACTGCGCGCAATTGG 877
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1141 rThrHisProLeuGlyLysAspValValGlnProThrAspAsnGlnPhe 1158
878 TTGCACAGACGACCGGTGATTCCCGTTCGGTATTGAAGCGCGGATT 927
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1158 lAlleThrPheAsnAlaIleArgGlyLysValThrValLysAlaVal 1174
928 ACACA.....GCGCGCACGATTTATTTGGAGCGGTACAC... 963
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1175 ValGlnLeuAsnGlnLysAlaArgLysHisPheThrGlyArgThrSerTh 1191
963 ..... 963
1191 rPheThrAlaGlnThrThrIleThrValGlnAspGlyLeuSerLeuValG 1208
964 .....AATCAGATTTCGCTATCGAAGAACGCGCGAC 996
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1208 LnProluMetAspIleAsnThrValArgValAlaProAsnSerGlnLeu 1224
997 AAGAGCTGTTCGCTGCGTTCGCGCGCAG.....CCGACAAATA 1037
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1225 LysMetValThrAlaThrPheGlnAlaSerPheSerValProSerAspPh 1241
1038 CTCG.....ATCAGCGGTACACCCCTGGCCATTCTCTGAAAAACAAC 1081
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1241 eSerSerArgLysValIleSerAlaAspIleHisLeuIleThrAsnGly 1257
1082 TCTTCAGTTCACACAGCC.....GTCAACGCGCGC 1113
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1258 .....LysGlnIleSerAlaAlaIleThrValArgAsnValAsnSerPro 1272
1114 GACGC.....GCCATGTGCGCATTTGCTTACGAGCCGCGATGCC 1157
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1273 AspAsnGlnThrValLeuIleProVal...ThrValSerArgValAlaSe 1288
1158 CTTCGATATCTGCGCCACCTGCTTTGCGC 1188
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1288 rLeuAspValHisProThrIleGlnLeuLys 1298

```

seq_name: p1r2:H70108

seq_documentation_block:
 hypothetical protein BB0072 - Lyme disease spirochete
 C/Species: Borrelia burgdorferi (Lyme disease spirochete)
 C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
 C/Accession: H70108
 R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 A/Authors: Smith, H.O.; Venter, J.C.
 A/Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
 A/Accession: H70108
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-773 <K1E>
 A/Cross-references: GB:AE001120; GB:AE000783; NID:g2687951; PIDN:AC66467.1; PID:g268
 A/Experimental source: strain B31

alignment_scores:

Quality:	Ratio:	Length:
109.50	0.684	306
Percent Similarity: 52.288	Percent Identity: 21.569	

alignment_block:

US-09-303-518d-125 x H70108 ..
 Align seg 1/1 to: H70108 from: 1 to: 773

76 ATTAGCGAAGTCCGCTTGGCGGAAGATATGCGGATATGCGCCCTC 125

[illegible]

301 AenguiLysIleasPThr 306
seq_name: p1r2:T34369
seq_documentation_block:
hypothetical protein T19D12.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34369
R:Favella, A.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid T19D12.
A:Reference number: Z21513
A:Accession: T34369
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1777 <FAV>
A:Cross-references: EMBL:U041263; PTDN:A2C24428.1; GSPDB:GN00020; CESP:T19D12.1
A:Experimental source: strain Bristol N2; clone T19D12
C:Genetics:
A:Gene: CESP:T19D12.1
A:Map position: 2
A:Introns: 36/1; 134/2; 180/1; 622/3; 691/2; 754/1; 1111/2; 1174/1; 1271/3; 1322/2; 1

```

alignment_scores:      length: 317
  Quality: 109.50
  Ratio: 0.788
  Gaps: 14
Percent Similarity: 43.849  Percent Identity: 24.606

alignment_block:
US-09-303-518D-125/rev x T34369  ..

Align seg 1/1  to: T34369  from: 1  to: 1777

```

```

935 CCTGTGTATATGCGCCGCTCAATATACGAACCGGAATATCAGCGGTTC 886
      ||| ||| ::: ::::||||: |||
310 PrivaValThrThrSerThrSerThrGlnGlyIleSerThrThrAl 326
      : :::||| ||| ::::|||||
885 TGAGTCAACCAATTCGCCGAGTAAT... 858
      : :::||| ||| ::::|||||
326 aglnAlaThrProSerSerSerValIleProThrThrGlnThrG 343
857 .....TGCATACTTTCGACCCCAAAACGGTACCAAGAGCGCGGT 816
      :::||| |||:||||||| ::::||||
343 InArgProThrSerThrGlyIleProSerThrValSerThrSergInGly 359
815 TTGTACTTGAAGACCACTAGGCAATCAGCAACGCGGTCTGGGTTCACAG 766
360 ThrSerSerThrSerProIleProSerThrThrGlnThrSerSerAl 376
765 GCCTGTTCGAACAAACGCCAATGTAATTCATCTTGTAATTGATG 716
      |||
376 Apo.....Ser 379
715 TCCACAGGTTTATTCGCGCGACGCGCTCGATGAATGAATGATGCGTG 666
      ::::|||| |||:|||||: :::
379 hrrYThrSerAlaPheThrProSerProThrThrIleuLeuThrSer 395
665 CCACTCAAAACGCGACGATGCGGCGCGCAATTCATGCTGTTGATGTT 616
      ::: ||::|||: ||| |||:||||
396 ThrIleAlaProSerThrGlnGlyAlaProThrSerSerIysSer... 410
615 GCGACGATTTCAGAGCGACGACTTCGCGCAAGTGCCCTTACAACATGGA 566
411 .....SerSerProAlaSerThrProThr... 418
565 TTTTGGCTTCGTCAAACGCTCAATACCAACAGGCGCGTTGAAATTC 516
419 .....ThrThrIleThrProGlyAlaProSer 427
515 TCGCGGCGCTTGTGATATGACCTAGGCTCGGACGACCGGATGCT 466
      ||| |||:|||||: ::::||||

```

```

428 Ser.....ThrLeuGlySerSerSerThrIleLeu 438
465 GTCCATGCATGTCAGACAGATGCGACACCGCTCGCATCGACGACGAGA 416
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
438 lSerThrThrIleThrPro...SerThrProLysValSerThrIleThrL 454
    :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
415 TTTTGCTGAACGACGCGGTGCGACGCGATCCACAAACCGGATTGGATC 366
    :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
454 eu.....SerGlnSerProThrPro..... 460
365 AGATTGCGGCGCATCTTTCGCCGCTTAAGTTGGCAGGCGCTTCAGTGC 316
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
461 .....ThrSerThrProLeuValValSerSerSer..... 471
315 GTACGCTTCAACCTGATTTCGTGCTGCTTCAACGCGCAATCAGCAGTG 266
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
472 .....SerGlySerSerSerThrValValThrSerThrIleThrPro. 485
265 ACTGAAGTACGCGCTTTTCGACGAGTGAATCGCGCATTTTGCCTGAA 216
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
466 .....SerThrGlnGly.....ValProThr 492
215 GCGGCGCAGTAACACACACCGCGGATTTTTCCTTCAACACAGCAC 166
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
493 SerThrSerAsnGlnProThrProSer.....ThrSerAsnProThr 506
165 TTGGGCT.....TTTTCAGCGCATGCGCTTCCTTACCTTTCATCGAGG 122
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
506 rThrProLysSerThrValThrAlaSerProSerThrThrGlyAlaThrS 523
121 GCGCGCA.....CCGCAATATCTTCGCCACGACGCGACTGGGTATG 75
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
523 erThrAlaSerSerThrThrIleThrSerSerAlaProThrSerGlnSer 539
74 GCGGCGCGTGTAAACGCGCTTCTCGGTTCGCCGCGATGGCGCAGTT 25
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
540 HisSerProSerSerThrMetThrSerThrValProValThrSerThrph 556
24 T 24
556 e 556

```

seq_name: pir2.T34434

seq_documentation_block:
 hypothetical protein K06A9.1a - Caenorhabditis elegans
 C.Species: Caenorhabditis elegans
 C.Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
 C.Accession: T34434
 R.Gelisel, C.: Gallung, S.
 Submitted to the EMBL Data Library, December 1996
 A.Description: The sequence of C. elegans cosmid K06A9.
 A.Reference number: 221525
 A.Accession: T34434
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-2232 <GEI>
 A.Cross-references: EMBL:U08046; PIDN:AC070890.1; GSPDB:GN00028; CESP:K06A9.1a
 A.Experimental source: strain Bristol N2; clone K06A9
 C.Genetics:
 A.Gene: CESP:K06A9.1a
 A.Map position: X
 A.Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/1

alignment_scores:
 Quality: 109.50 Length: 458
 Ratio: 0.495 Gaps: 18
 Percent Similarity: 48.253 Percent Identity: 22.926

alignment_block:
 US-09-303-518D-125/rev x T34434

```

Align seg 1/1 to: T34434 from: 1 to: 2232
1340 CTTTCCTTCATAGTGTTCACGACTTTCGCAACAGCGGCGGTATTC 1291
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1425 ProThrGlySerThrGlnSerSerThrIleLeuSer.....Se 1437
1290 GTATTTCGCGGCGACGACGAGCTGCACAAAGCAGAGTCTTTCGTCA 1241
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1437 rThrIleSerGlySerThrGln...HisThrIleMetSerIysAlaSerS 1453
1240 ATTCCAAACCAACCAATGCTGCGCGCTTCGGTATCGCGCAGATTA 1191
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1453 ergly.....SerThrProSerThrAsn 1461
1190 TCGCGCAAAAGCAGGCTGGCGAGATATCCAAAGGCGATCAGCGCTG 1141
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1462 SerGln.....ThrGlySerThr 1467
1140 AGTACCAATGCGACCATGCGCGGCTGCGCGCTTGACGCGTGTGA 1091
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1467 rValThrMetGlySerSerSerThrSerGlyValSerThr..... 1480
1090 ACTTGAAGATTGTTTTCAGGAAATGCGCGAGGTTGTACGCGTATG 1041
1480 ..... 1480
1040 GATATTTCGCGCGCTGCGCGGCGCAACCGCAACGCTTTCCTGCG 991
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1481 .....SerSerAlaSerSerThrGlnProGlnMetSerThrSergl 1494
990 GCGTTTCGATACGAAATCTGATTGCGTACGCTCCCAATTAATGT 941
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
1494 nGlySerSerAlaGlySer.....ThrV 1502
940 GCGGCGCTGTGTAATCGGCGG.....TTCATACCGAACCGGAATC 897
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1502 aAlaSerSerThrAlaSerProAlaIaSerSerThrAlaProSerSer 1518
896 ACGCGTGTCTGTGTCACCAACCAATGCGCGCGAGTA.....ATTG 856
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1519 ThrGlyThrMetSerSerThrSerSerGlyThrValGlySerThrIleSe 1535
855 CGATCTTTCGACCCCAACGTAACGACGACGAGCGGCTTGTGACTT 806
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
1535 rGlnSerSerThrThrAlaSerAlaSerSerGlnThrGly...SerThrV 1551
805 GAGAACCACTAGGCGCATCAGCGGCTGCTTCACAGCGCTGTGCA 756
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
1551 alThrMetGlySerSerSerSerThrSerGlyValSerThrSerSerAlaSer 1567
755 AACCAACGCGCAATGATTAATACATCTTGATTAATGATGTC...CAC 709
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
1568 SerThrGlnProGlnMetSerThrSerGlnGlySerSerAlaGlySerTh 1584
708 GGTTTTATTCGCGCCGACCGGCTCGATGAATGAATGCGTGCACCTCA 659
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1564 rValAlaSerSerThrAlaGlyLeuValSerThrSerThrValProSerS 1601
658 AACCGGCGATGCGGCGCGCGCAATTCATGCTTTGATGATGTCGACGA 609
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
1601 erThrGlyThrMetGlySerThrSerSerGlyThrValGlySerThrIle 1617
608 TTTTCAGACGCGACGCTGCGCGACGCTTTCACAAACATGATTTTG... 561
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
1618 SerGlnSerSerThrThrAlaSerAlaSerSerGlnThrGlySerThrVa 1634
560 .....CGTTGCTCAACGCGCTCAATA 539
1634 lThrMetGlySerSerSerThrSerGlyValSerThrSerSerAlaSerS 1651
538 CCAACAGCGCGCTTGAATCTCG...GCGGCTCTTTGATTAATGACG 492
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
1651 erThrGlnProGlnMetSerThrSerGlnGlySerSerAlaGlySerThr 1667

```


[illegible]

```

seq_documentation_block:
  H+/K+-exchanging ATPase (EC 3.6.1.36) [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: A980048
R:Holshuis, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burger, S.; Dehoff, B.S.;
  e, R.; Leblanc, D.J.; Lee, L.N.; Letkowitz, E.J.; Lu, J.; Matsushima, P.; McAnen, S.;
  y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.
A:title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A980048
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-898 <RUR>
A:Cross-references: GB:AE007317; PIDN:AAU00214.1; PTD:915459062; GSPDB:GN00174
C:Genetics:
A:Gene: pact
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: hydrolase

```

```

Ratio: 0.617 Gaps: 15
Percent Similarity: 49.157 Percent Identity: 23.315
alignment block:
US-09-303-518D-125 x A98048 ..
Align seg 1/1 to: A98048 from: 1 to: 898

```

28	CTGGCCATTCGGGGGACAGCCGACCAAGCCGTTTACGACGGCCGGCCAT	77
287	LeuAlaValAlaAlaAlaIleProGluGlyLeu.....ProAlaI	299
78	TACCGAAGTCGCGTTG...CTGGCGGAAGATATCGCGTATGCGCCCT	124
299	evalThrIleValIleuSerLeuGlyThrGlnValIleuAlaLysArgHis	310
125	CGATG.....AAGTCAAGGAAGCGCATGGCCGTCAAAAAAGCCAAATG	166
316	erIleValAlaGlyLysLeuProAlaValGluThrLeuGlySerThrGluIle	333
169	CTGTTTGAGACAAAGAAAGAT.....CC	191
333	IleAlaSerAspLysThrGlnGlyThrLeuThrMetAsnLysMetThrValG	349
192	GGCGCGTGTCTTACTGCGCGCGCTTACGCAAAATGCGCGCATTCACC	241
349	uLysValIlePheTyrAspAlaValIleuHisAspSerAlaAspArgIle	366
242	GTCGC...GAAGACGGGTACTTCAGTACATCGTGAATGCGCGTTGAAGC	288
366	euGlyLeuGluMetProIleuLeuArgSerValIleuAlaAsnAspThr	382
289	AACGACGAATCGATTTTGAAACGCTACGACCTGAAAGCGCTGGCAACT	338
383	LysIleAspValGlu.....GlyAsnLe	390
339	AAGCGCGGAAGAGTCGCGCGCACTGATTCATCGCGTGTGTGAGTCG	398
390	uIleGlyAspProThrGluThrAlaPheIleGlnTyrAlaIleAspLysG	407
389	CGTCGGCACCGCGTCGTTG...ACGAAATTCGTCGCGTACGCGGAG	435
407	LTTyrAspValLysGlyPheLeuGlnLysTyrProArgValAlaGluLeu	423
436	CCGTGCGCATTTGCTGATCGATCGATGAGCAAC...AATCGCGTGGTCG	482
424	ProPheAspSerAspArgLysLeuMetSerThrValHisProLeuProAs	440
483	GCACCCCTACGCTATTATCAAGAAAGCGCGCGCAGATTTCAAACGCGGCC	532
440	pSerArgPheLeuValAlaValLysGlyAlaProAsp.....	452
533	TGTTGGTATGACCGGTTTGACCGGAAGCGAAATGCAATGTTGTTAAGCA	582
453GlnLeuLeuLysArgCysLeuLeuArgAspLysAla	464
553	GCTGGCGAGACGTCGCGTGTGAAGATGCTGGCAAC...ATCGAAACA	629
465	GlyAspIleAlaProIleAspArgLysValThrAsnLeuIleHisThrAs	481
650	TGAATTCGGCGGCGCCGCACTCTGCC.....GGTTTGATGCGCAGCACA	673
481	nasSerGlnMetAlaHisGlnAlaIleuArgValIleuAlaGlyIleTyr	497
674	TTCATTATGACGAGCGGCTCGCGCGGATTAACACCGTGTGACATCATAT	723
498	..LysIleIleAspSerIleProGluAsnLeuThrSerGluGluIleuGlu	513
724	TATCAAGATGAATATACATTCGTCGCGGTTTGTTCGACAGCGCCGCTGAA	773
514	...AsnAspLeuIlePheThrGlyLeuIle.....GlyMetIleAs	526
774	CACCGAGCGCGGTATGTGCCCTAGGTGTTTCAAGTACAACAACCGCGCC	823

```

526 : ||||| ||| : : ||| |||
      ProgluArgProgluAlaAlaValArgValAlaLysGluLac 543
824 TCTTGCCTACCGTTTGGTGGCGAAGATATCGCAATCTACTGGGGGAA 873
      : ||| : : ||| |||
543 LylleArgProLleMetleThrcLysPheSphLysThrAlaGluAla 559
874 TTGGTT ..... GACACAGACAACCGCT 896
      : : : : :
560 lLeAlaLysArgLeuGlylleleAspAlaAsnAspThrGluGlyHisVa 576
897 GATTTCGGCTGCTATTGAACGCGCGATTAACAGCGCGCGATTT-946
      : : : : : ||||| : : : : :
576 lLeThrcLylAlaGluLeuAsnGluLeuSerAspGluGluLysV 593
947 ATTTGGAGCGCTACCAATCATGATTCCGTTATCGAAGAGCGCGAGC 996
      : : : : : ||| : : : : :
593 alValGlyLntYrSerValTyrAlaArgValSerProGluHisLysVal 609
997 AAGAGCGTTCGGCTGG 1014
      : : : : : |||
610 ArgLeuAlaLysAlaTrr 615

```

seq_name: p1r1:A36904

seq_documentation_block:

Prochlorophyllide reductase (EC 1.3.1.33) [similarity] - *Synechococcus* sp. (PCC 7942)

C:Species: *Synechococcus* sp.

A:Variety: PCC 7942

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: S25621; A36904

R:Item: Hurwitz, J.; Ronen-Terazi, M.; Gabai, C.; Hassidim, M.; Schwarz, R.; Kaplan, A. submitted to the EMBL Data Library, August 1992

A:Reference number: S25621

A:Accession: A36904

A:Molecule type: DNA

A:Residues: 1-466 <LIE>

A:Cross-references: EMBL:X67694; NID:9312508; PIDN:CAA47924.1; PID:g46847

R:Price, G.D.; Howitt, S.M.; Harrison, K.; Badger, M.R.

J. Bacteriol. 175, 2871-2879, 1993

A:Title: Analysis of a genomic DNA region from the cyanobacterium *Synechococcus* sp. str.

A:Reference number: A36904; M0ID:93259930

A:Accession: A36904

A:Molecule type: DNA

A:Residues: 388-466 <PRI>

A:Experimental source: PCC 7942

A:Note: Sequence extracted from NCBI backbone (NCBIN:131726, NCBI:P:131727)

C:Superfamily: prochlorophyllide reductase chain chlN

C:Keywords: chlorophyll biosynthesis; oxidoreductase; photosynthesis

alignment_scores:

Quality: 107.00

Length: 333

Ratio: 0.728

Gaps: 16

Percent similarity: 44.144

Percent identity: 21.021

alignment_block:

US-09-303-518d-125 x A36904 ..

Align seg 1/1 to: A36904 from: 1 to: 466

```

7 AAAAAACAAAAAGCTAAACCTGACCGATCGGCGGACAGCGGAGCAAGC 56
  : ||||| : : |||
92 GlnlleYsAlaGAspArgAsn.....ProSeValIlI 102
57 CATTACAGAGCGCGCGCATTAACGAGATCGCGTTGCTGGCGAAGAT 106
  : ||| : : ||||| : : :
102 eValTrpIleGlyThrCysThrThrGluIlelleYsMetAspLeuGlu 118
107 ATGCCGATATCGCGCGCTTCGATGAAGATCAAGAGAGCGGATCCCTCAA 156
  : ||| : : ||||| : : |||
119 .....GlyLeuAlaProYsleuGluAlaGluIleGlyllePro..... 131

```

```

157 AAGGCGCAAGTCTGTTTGAAGACAAAAAAGATCCGGCGCTGCTTAC 206
      : : : : :
132 .....lleValVa 134
207 TGCGCGCGCTTACGCAAAATGCGCGATTCACCTGGCGGAAAGCGG 256
      : ||| : : ||||| : : |||
134 lAlaArgAlaAsnGlyLeuAspTyrAlaPheThrcGlnGlyAspTrhV 151
257 TACTTCAGTCAGTCGATGATGGCGCTTGAAGGCAACGACCAATTCAGTTT 306
      : : : : :
151 alLeuAlaAlaMetAlaAlaArg..... 158
307 GAACGCTACGACCTCAAGCGCTGGCAACTTAACGCGGAGAGATGCG 356
      : : ||||| : : : : : ||
159 .....CysProGluAlaAlaThrSerGluAlaAspGlnGluAlr 172
357 CCGCAACCTGATTCAAATCCGGTTTGTGACGCTGCGCACCGCTCGT 406
      : ||| : : ||||| : : |||
172 gThrAsnAlaIleGln.....ArgLeuGluGlnP 182
407 TCAGCAAAATTCCTGCGCTGATGCGGAG..... 435
182 heGlyLysSerProAlaAlaGluGlnGlnProAlaSerIlyshSpro 198
436 CCGTTCGCGCATCTGCTCAATGCGATGACACCAATCCGCTGCTGCCA 485
      : ||| : : ||||| : : |||
199 ProleuIleuPheGlySerValProAsp.....ProValAlaThrcL 213
486 CCTACGCTCATTTATCAAGAGAGCGCGGAGATTTCAACGCGGCTGT 535
      : : ||| : : ||||| : : |||
213 nleuThrIle.....GluLeuAlaLysGlnGlyIleT 224
536 TGCTATTGAGCGCTTGACCGAAGCAAAATTCATGTTGTAAGGAGCT 585
      : ||| : : ||||| : : |||
224 hValSerGlyTrpLeuProAlaLysArgTyr..... 234
586 GCGCAGAGCGTGCGC.....TCTGAAATGCTGCCAATCATCAACACA 629
      : : ||| : : ||||| : : |||
235 ...ThrGluLeuProValIleAlaGlnGlySerTyrAlaIleGlyLeuAs 250
630 TGAATTC..... 636
250 nProPheLeuSerArgThrAlaThrThrLeuMetArgArgLysCysL 267
637 .....GGCGCGCGCATCTGCGGCTTGTGAGCGACGAC..... 672
267 ysValIleGlyAlaProPheProIleGlyProAspLysSerArgAlaTrp 283
673 .....ATTCAATTCATGACCGCGTGGC..... 696
284 lLeGluLysIleCysSerValleuGluIleGluProGlnGlyLeuAlaG 300
697 .GCGAATAAACCGTGTGACCATC.....AATTATCAAGATGTAATTA 739
      : : ||| : : ||||| : : |||
300 uArgGluAlaGlnValTrpAspSerIleGluAspTyrArgGlnLeuValG 317
740 CCAATGCGCGTTGTTGTCACACAGC.....GCTGACACAC 777
      : : ||| : : ||||| : : |||
317 lueGlyLysGlnValPhePheMetGlyAspAsnLeuTrpGluIleSerLeu 333
778 GAGCGCGATGATGCGCTAGTGT.....TCTCAAGTCAACAAC 818
      : ||| : : ||||| : : |||
334 AlaArgPheLeuValArgCysGlyMetArgCysProGluIleGlyLeP 350
819 GCGCTCTTGCCTACCGCTTTGGTGGCGAAGATATCGCAATTAACGCG 867
      : ||| : : ||||| : : |||
350 oTyrLeuAspArgArgTyrLeuGlyAlaGluLeuAlaMetLeuGluAla 366

```

seq_name: p1r2:E95180

seq_documentation_block:

cation-transporting ATPase, E1-E2 family SPI51 [imported] - *Streptococcus pneumoniae*

C:Species: *Streptococcus pneumoniae*

C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C/Accession: E95180
R/Tellelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: E95180
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-914 <KDR>
A:Cross-references: GB:AE005672; PIDN:AAK75638.1; PID:g14973040; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
A:Gene: SPI551
C:Superfamily: Nat/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain

alignment_scores:
Quality: 107.00 Length: 356
Ratio: 0.615 Gaps: 15
Percent Similarity: 48.876 Percent Identity: 23.315

alignment_block:
US-09-303-518d-125 x E95180 ..

Align seg 1/1 to: E95180 from: 1 to: 914

```
28 CTGCCCATCGCGGCGGAGCCGAGCAACCGCTTACGACGCCCGCCCAT 77
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
303 LeuAlaValAlaAlaAlaLeuProGluLeu.....ProAlaI 315
78 TACCGAATCGCGCTTG...CTTGGGAGAGATATGCGGTATGCGCCCT 124
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
315 eValhrlleValleuSerleuGlyThrGlnValleuAlaLysArgHis 332
125 CGATG.....AAAGTCAGAGAGCGGATGCGGTCAAAAAGGCCAAGT 168
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
332 erlleValArgLysleuProAlaValGluThrleuGlySerThrGluIle 348
169 CTGTTTGAAGACAAAAGAAAT.....CC 191
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
349 lleAlaSerAspLysThrGlyThrleuThrMetAsnLysMetThrVal 365
192 GGGCGTGTGTTTACTGCGCGGCTTCAGCAAAATCGCGCATTCACC 241
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
365 ulysValPheThrAspAlaValleuHisAspSerAlaAspAspIleGlu 382
242 GTGGC...GAAAAGCGGTACTGATGATGATGATGATGATGATGATG 288
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
382 euGlyLeuGluMetProleuLeuArgSerValValleuAlaAsnAspThr 398
289 AACGACGAATCGAGTTTGAACGCTACGACCTGAAAGCGCTGGCAACT 338
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
399 LyslleAspValGlu.....GlyAsnLe 406
339 AAGCGGAGAAAGTGGCGCGCACTGATCAATCGGTTTGTGACGTG 388
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
406 uilleGlyAspProThrGlnThrAlaPheIleGlnThrAlaLeuAspLys 423
389 CGGTGGCGACCGCGCTTC...AGCAAAATCTCGCGCTGATGCCGAG 435
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
423 llyThrAspValLysGlyPheleuGluLysThrProArgValAlaGluLeu 439
436 CCGTTCCCATCTTCGTCATCGATGCGATGACACC...AATCGCTGGCTG 482
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
440 ProheAspSerAspArgLysleuMetSerThrValHisProleuProAs 456
483 CGACCTACGGTCATTATCAAGAAGCGCGGAGATTTCAAACGGCGCC 532
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
456 pGlyArgPheleuValAlaValLysGlyAlaProAsp..... 468
```

```
533 TGTGGTATGAGCGGTTTGACCGAAGCAAAATCCATGTTTGAAGCA 582
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
469 .....GlnleuLeuLysArgCysleuLeuArgAspLysAla 480
583 GCTGGCGCAGACGTGCCGTGAAAATGTCGCCAAC...ATCGAAACACA 629
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
481 GlyAspIleAlaProIleAspGluLysValThrAsnleuIleArgThrAs 497
630 TGAATTGGCGGCGCGCATCTGCC.....GATTGAGTGCGACGACCA 673
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
497 nasSerGluMetAlaHisGlnAlaLeuArgValleuAlaGlyAlaLys 513
674 TTCAATTTCATCGACCGCGTCGCGGCGGATTAACCGGTGGACCATCAT 723
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
514 ..LysIlelleAspSerIleProGluAsnleuThrSerGluGluLeuGln 529
724 TATCAAGATGATTAATTCATTTGCGCGTGTGTTGCAACAGCGCGTGA 773
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
530 ...AsnAspIleuIlePheThrGlyLeuIle.....GlyMetIleAs 542
774 CACCGAGCGCGGTATGCCCTAGGTGTTCTCAAGTCAACAAACGCGCC 823
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
542 pProGluArgProGluAlaAlaGluAlaValArgValAlaLysGluAla 559
824 TCTTCGCTACCGTTTGGGTCGGAAGTATCCGAATTTCTCGGCGGCA 873
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
559 llyleArgProIleMetIleThrGlyAspHisGlnAspThrAlaGluAla 575
874 TTGGTT.....GACACAGACACCGCGT 896
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
576 lleAlaLysArgleuGlylleIleAspAlaAsnAspThrGluGlyHisVa 592
897 GATTTCGCGTGGTATTGACGCGCGGATTACACAGCGCGCACAGATT 946
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
592 lleuThrGlyAlaGluLeuAsnGluLeuSerAspGluGluPheGluLys 609
947 ATTGGGAGCGTACACCAATCAGATTTCGTTATGAGAACGCGCGGACG 996
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
609 alValGlyGlnThrSerValThrAlaArgValSerProGluHisLysVal 625
997 AAGAGCTGTTCGGCTGG 1014
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
626 ArgIleValLysAlaLys 631
```